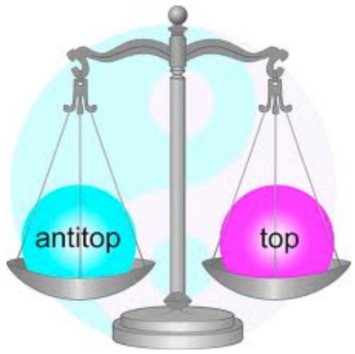
Measurement of the top quark pair production cross section at 7 TeV using secondary vertex b-tagging

Ioana Anghel

University of Illinois at Chicago



August 16, 2011

Outline

- Motivation
- Experiment Description
- Top Quark Pair Production at LHC
- Event Selection
- Top Pair Background Estimation
- b-tagging
- Results & Conclusions

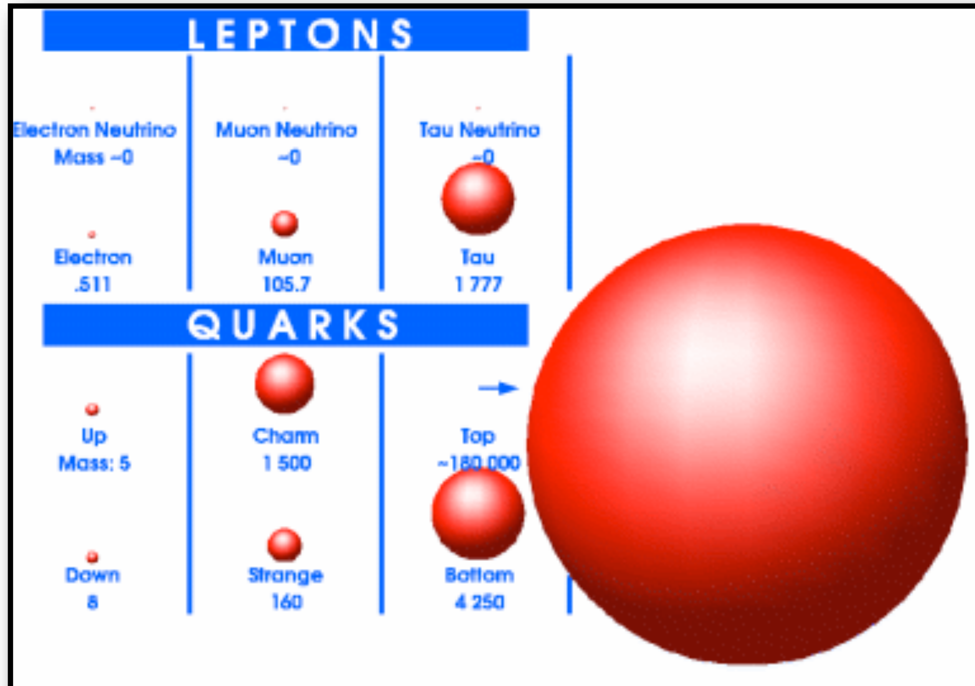
Top Quark Properties

- Heaviest fundamental particle

$m_t = 173.1 \pm 0.6(stat) \pm 1.1(syst)$ *Tevatron*

- Lifetime: 5×10^{-25} s

shorter than hadronization time, decays as a free quark



- Test of pQCD at high Q^2

Why do we measure Top pair production cross section ?

- Provides sample composition for other top properties measurements (charge, spin, SM electroweak interactions, coupling to particles)

- Gives input for searches for which top events are a dominant background

- Sensitive to new physics - Expect higher x-sec if non-SM production occurs

Top Quark Pair Production Cross Section

How do we measure the Top pair production cross section ?

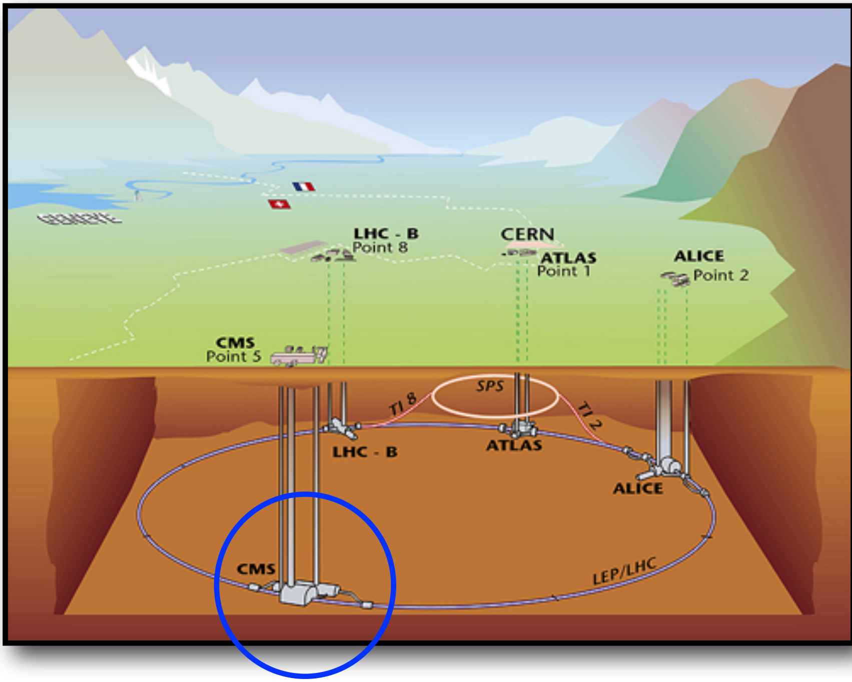
- Measure in different channels
- Measure with different techniques
  - with b-tagging (b-tagging method assumes Br(t→Wb)=1)
  - without b-tagging (kinematic fit methods are free of this assumption)
- This analysis: count the excess of events above background, after applying b-tagging

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N_{bkg}}{BR \times L \times \epsilon^{pres} \times P^{tag}}$$

where BR - the branching ratio of the final state,
L - the integrated luminosity,
ϵ^{pres} - the ttbar preselection efficiency,
P^{tag} - the probability of a ttbar event to have one or more jets identified as b jets.

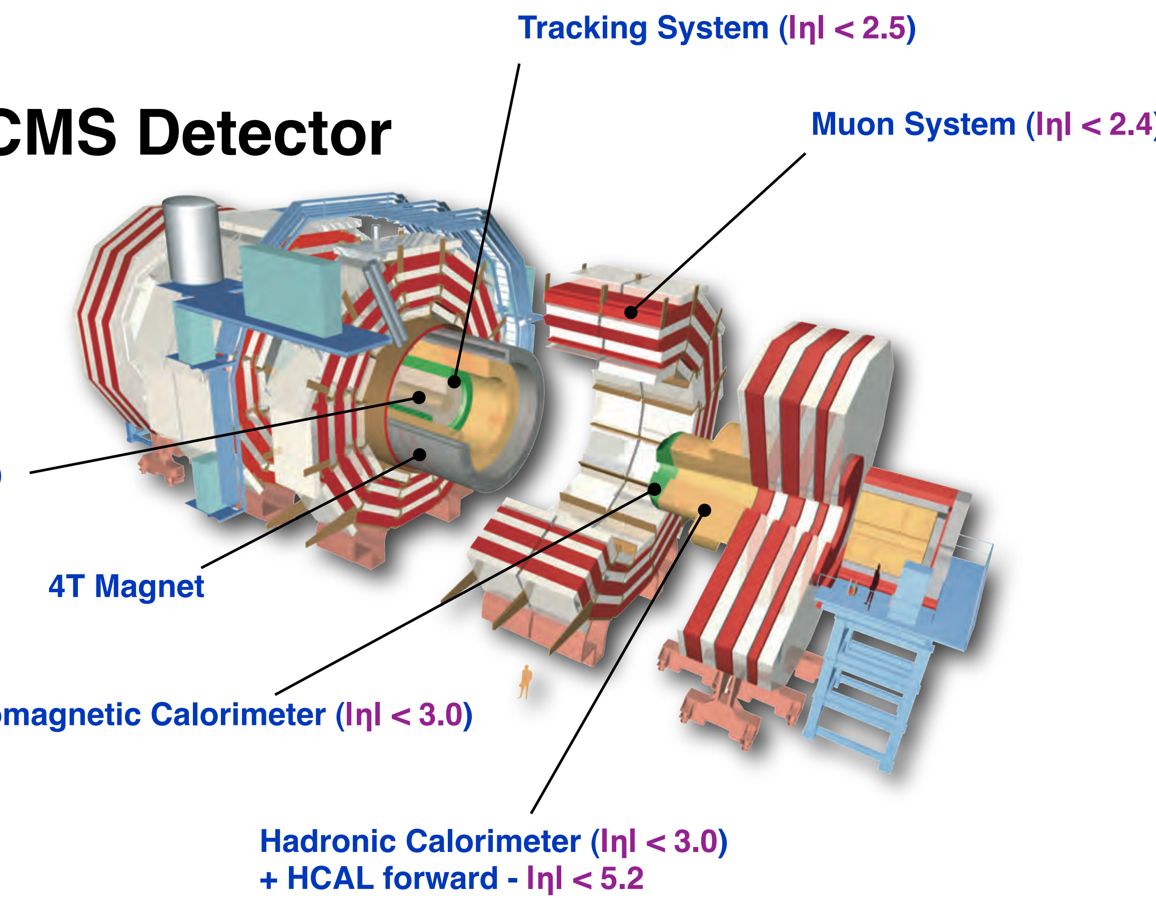
LHC and CMS

LHC



- Large Hadron Collider (LHC) - runs at 7 TeV center of mass energy since 2010

CMS Detector



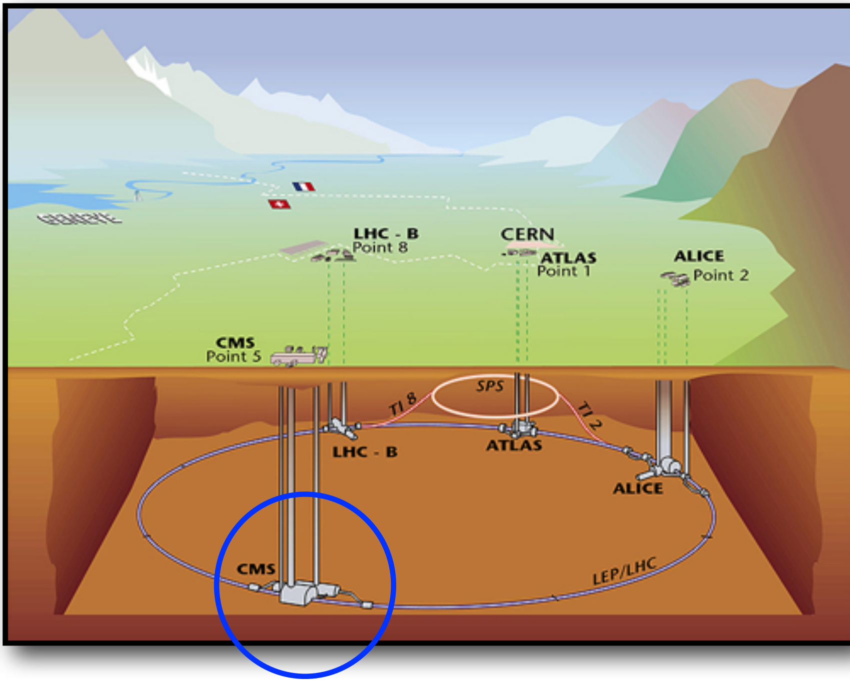
- CMS (Compact Muon Solenoid) characteristics:

  - CMS Tracker has the largest silicon area ever built, providing a very good resolution - helping in identifying the b-quarks

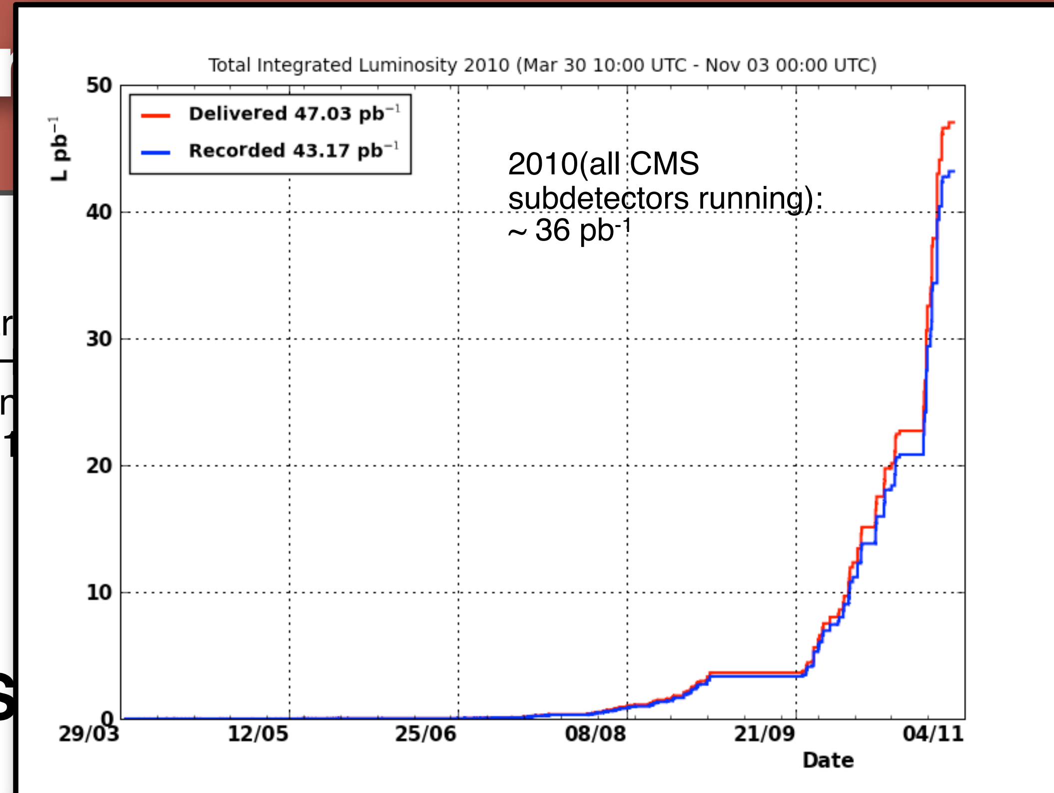
  - excellent instrument in muon identification

LHC ar

LHC



- Lar
(LH
cen
201



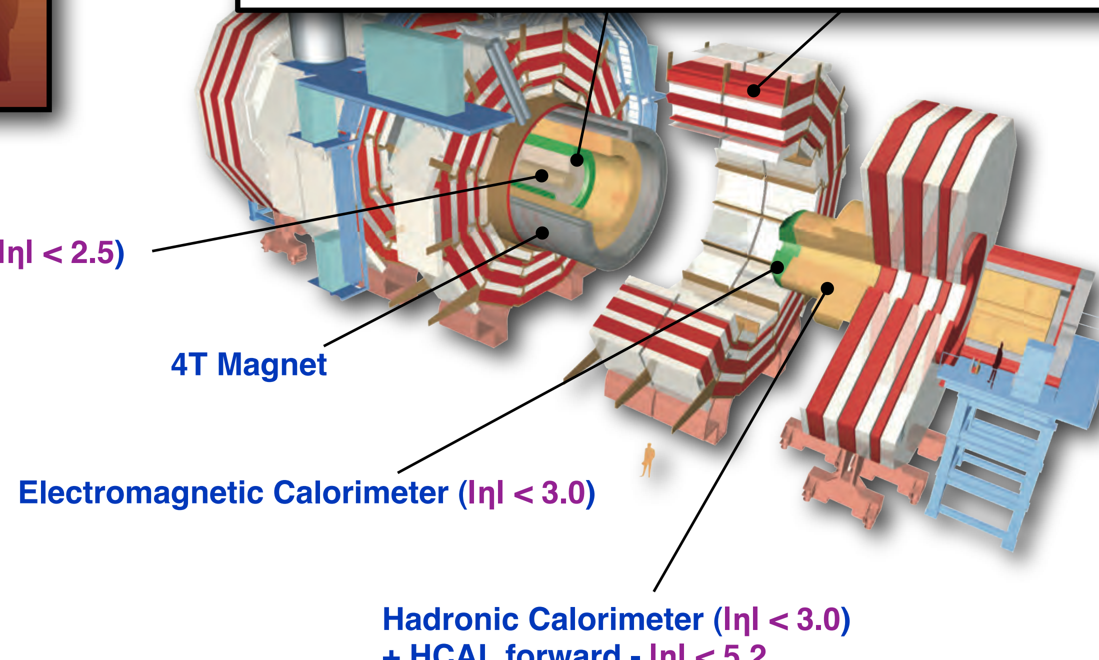
CMS

- CMS (Compact Muon Solenoid) characteristics:

  - CMS Tracker has the largest silicon area ever built, providing a very good resolution - helping in identifying the b-quarks

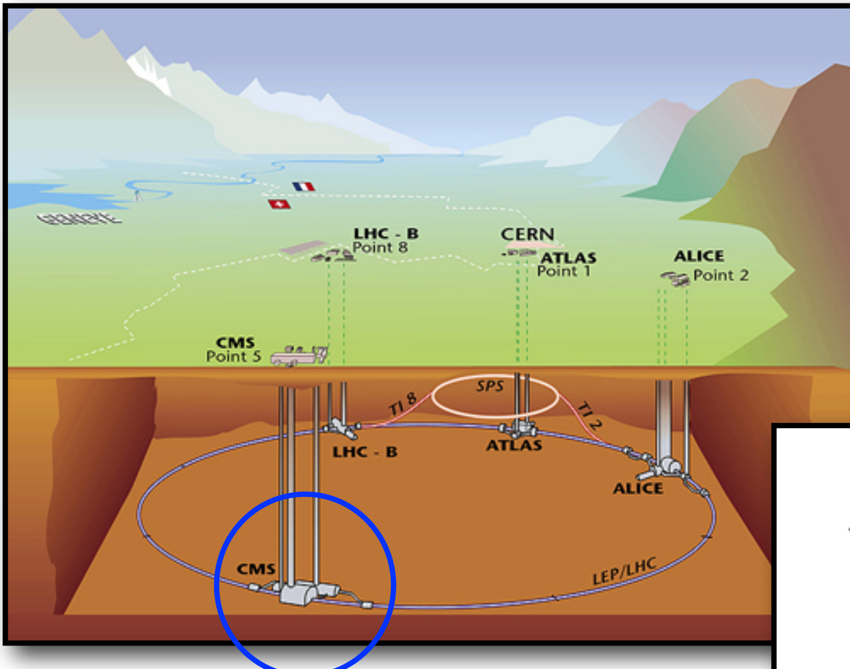
  - excellent instrument in muon identification

Pixel (|η| < 2.5)

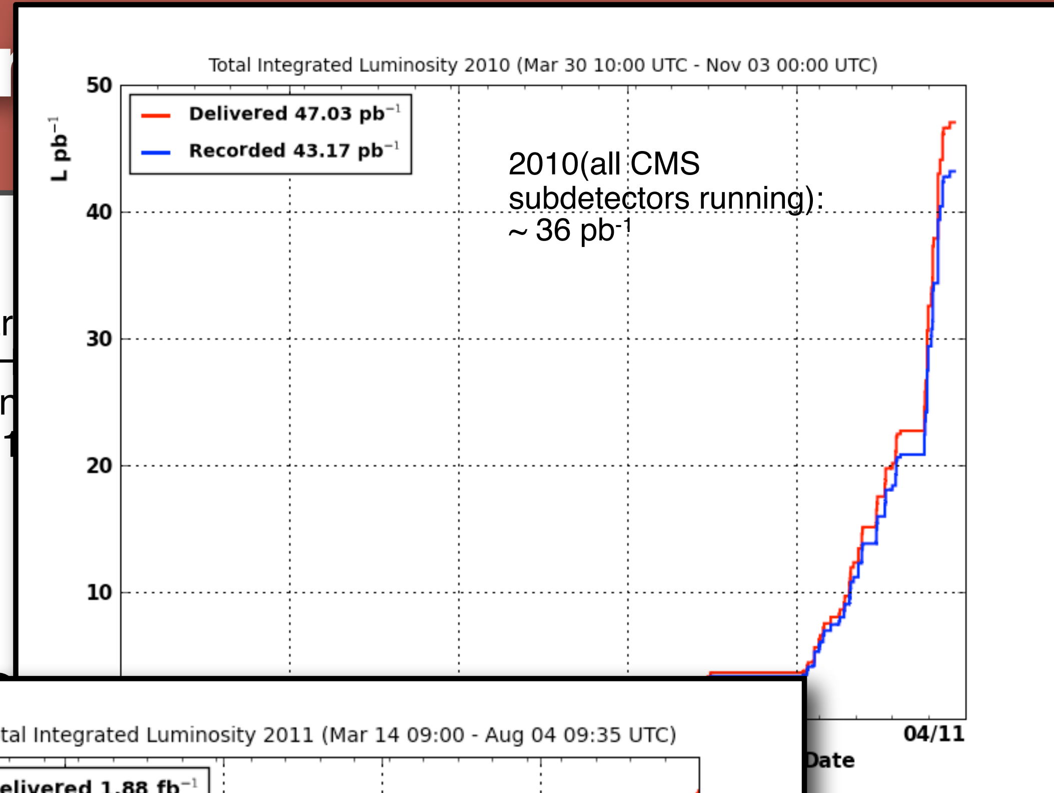


LHC ar

LHC



- Lar
(LH
cer
201

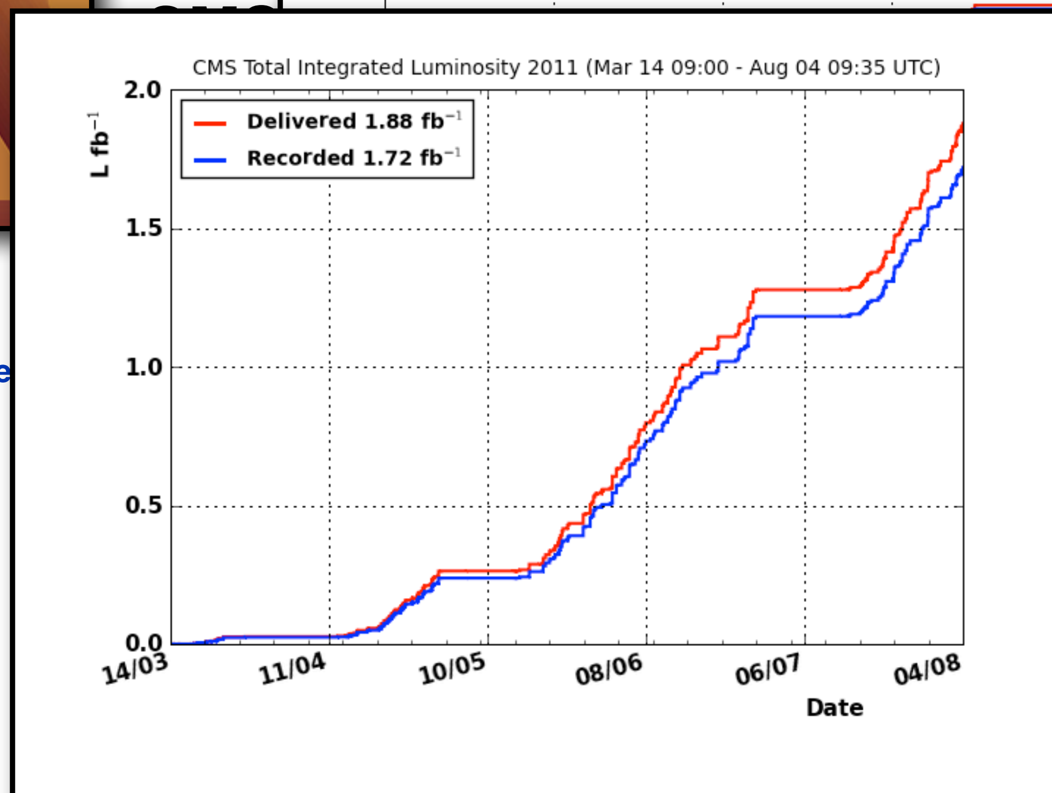


- CMS (Compact Muon Solenoid) characteristics:

Pixe

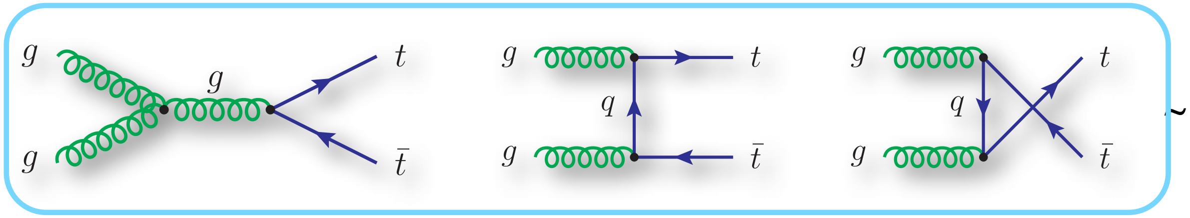
- CMS Tracker has the largest silicon area ever built, providing a very good resolution - helping in identifying the b-quarks

- excellent instrument in muon identification

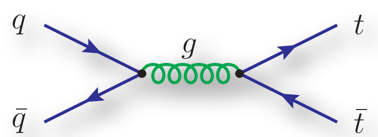


Top Quark Pair Production

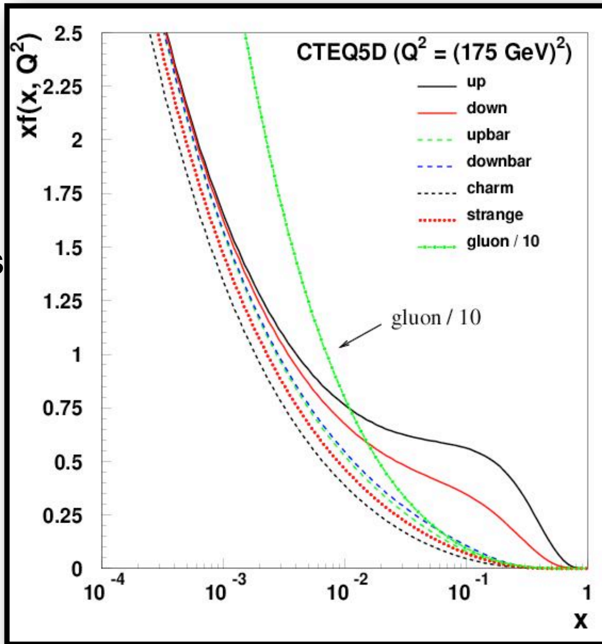
- At LHC, most of the top quarks are produced as $t\bar{t}$ pairs via strong interactions



~ 90% of processes

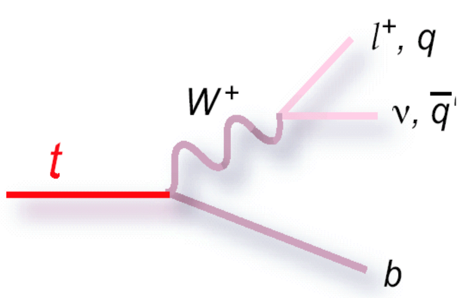


~ 10% of processes

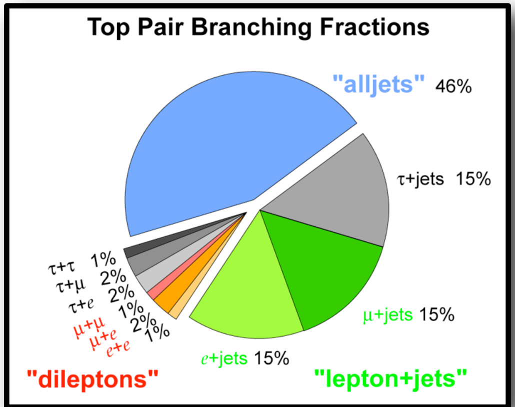


- Typically, sea and/or gluon interactions at low x dominate top quark pair production at LHC

- Within the Standard Model, the Top quark decays via the weak interaction exclusively as $t \to \mathrm{W}b$



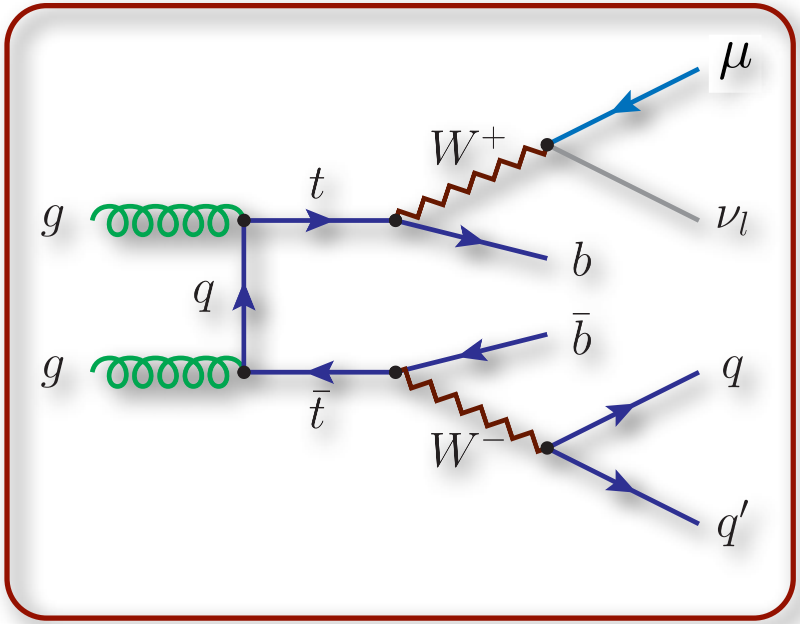
- The final state is determined by the decay of the W boson
  - dilepton
  - semi-leptonic
  - hadronic



* Tau's are treated separately due to their decay

Muon+Jets Channel

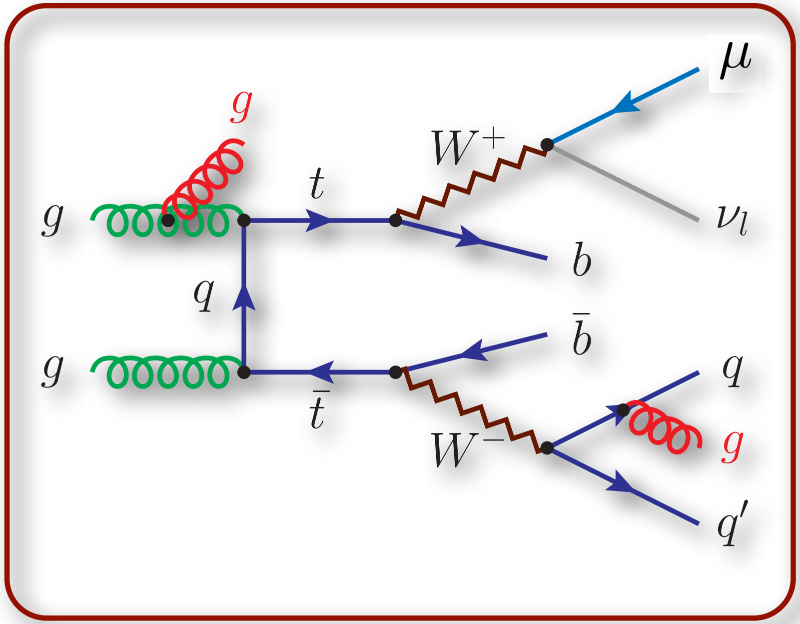
SIGNAL



- One isolated high p_T muon
- One energetic neutrino (MET)
- 4 high p_T jets

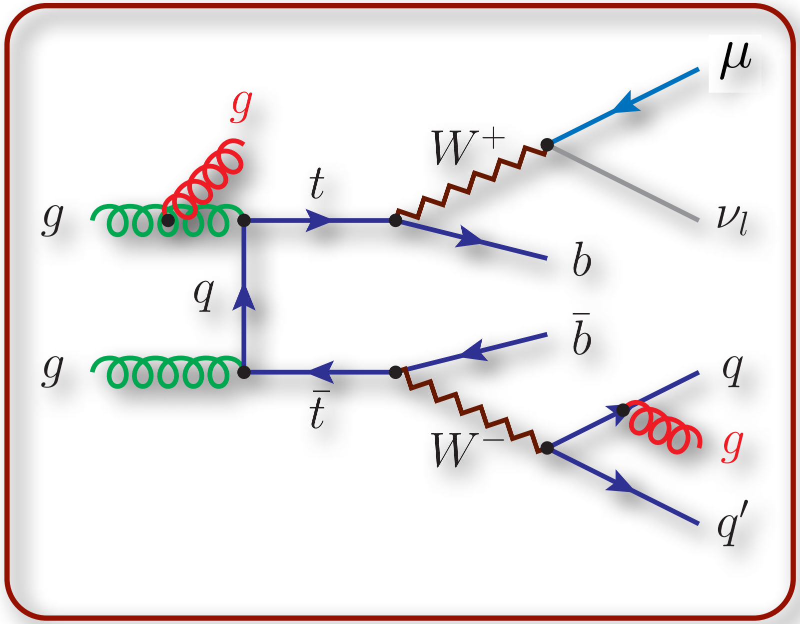
Muon+Jets Channel

SIGNAL



- One isolated high p_T muon
- One energetic neutrino (MET)
- \+ 4 high p_T jets

SIGNAL

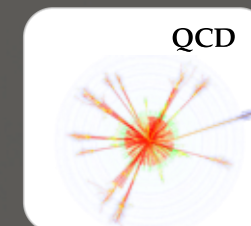


- One isolated high p_T muon
- One energetic neutrino (MET)
- \+ 4 high p_T jets

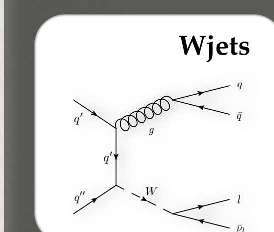
BACKGROUND

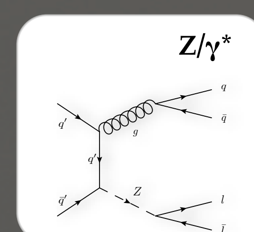
Instrumental

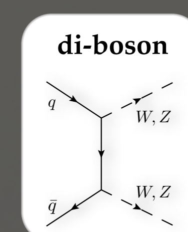
Jet mis-reconstruction or heavy flavor quark leptonic decay

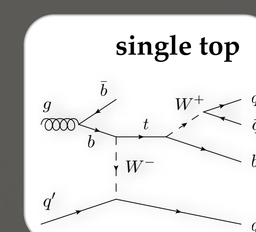


Physics









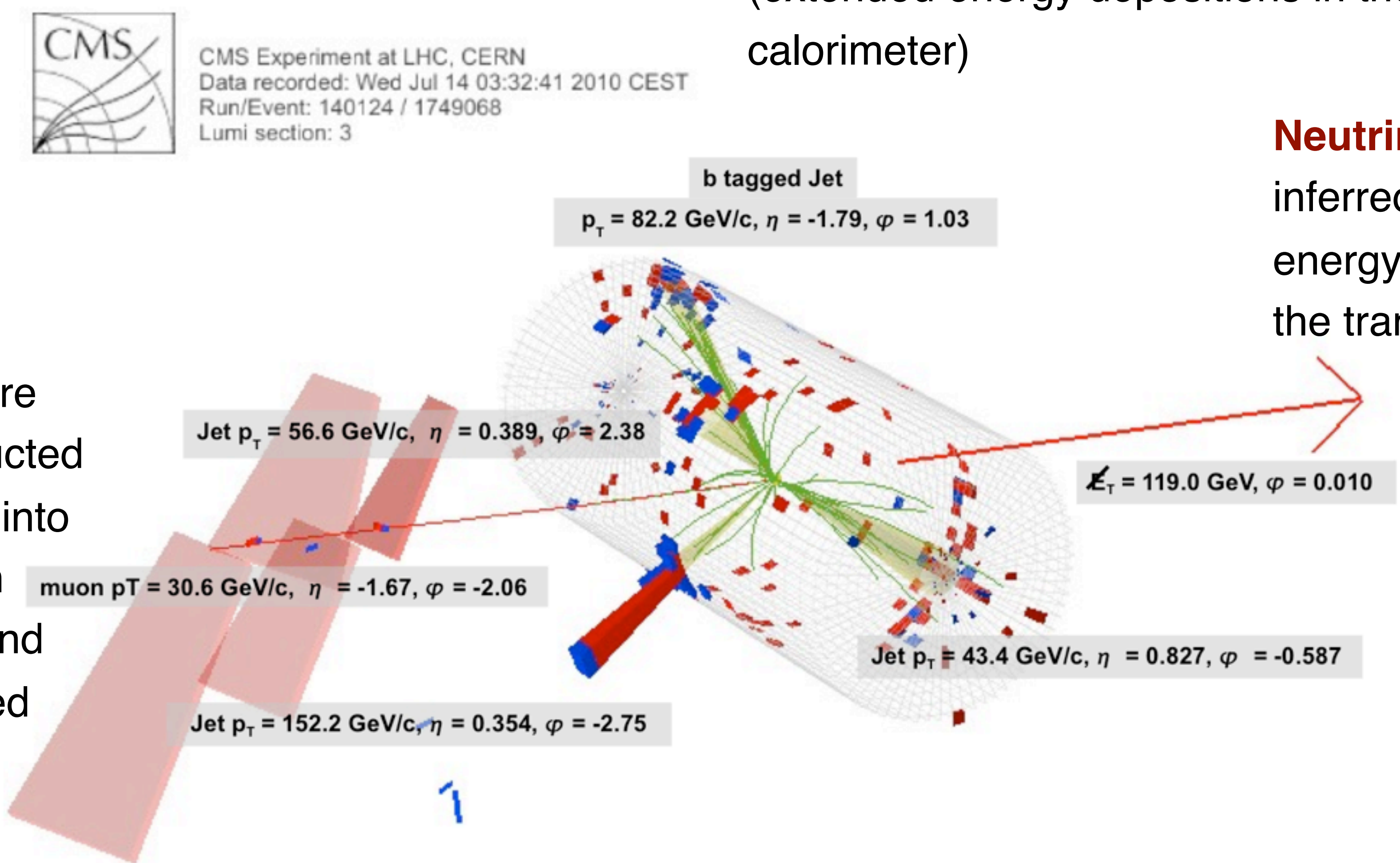
Always present and natural due to physics processes

Top Pair Event Candidate

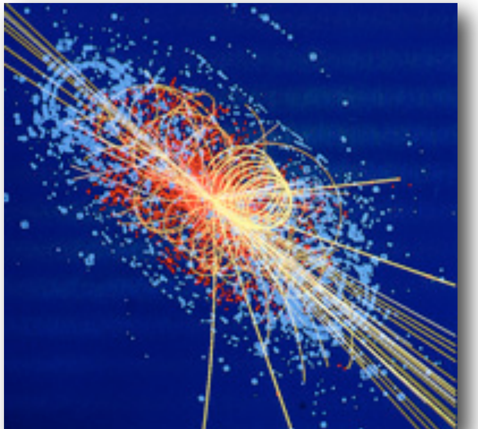
Quarks hadronize into **Jets** of particles (extended energy depositions in the hadronic calorimeter)

Neutrinos are inferred from the energy balance in the transverse plane

Muons are reconstructed from hits into the muon system and associated tracks

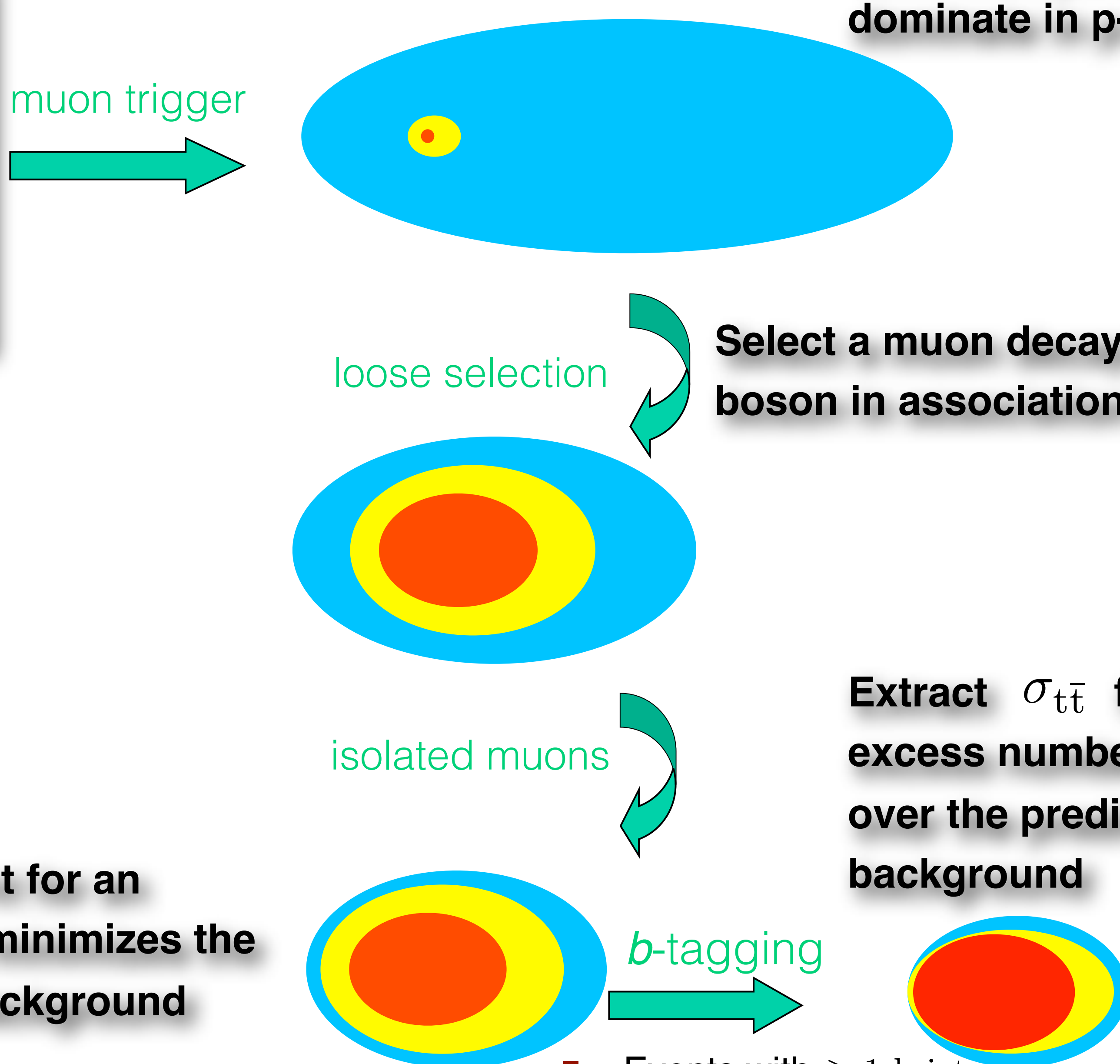


Event Selection



ttbar
physics bkg
QCD

QCD multijet events dominate in p-p collisions



Select a muon decaying W boson in association with jets

Extract $\sigma_{t\bar{t}}$ from the excess number of events over the predicted background

The requirement for an isolated muon minimizes the instrumental background

- Events with ≥ 1 b-jet
  ($t\bar{t}$ has 2 b-quarks!)

Event Selection (I)

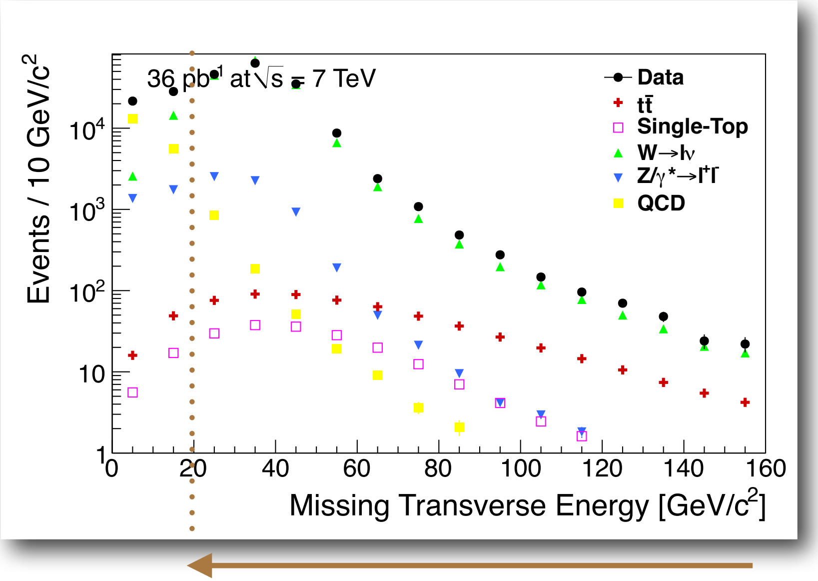
- Muon Trigger, exactly one Primary Vertex
- Missing transverse energy > 20 GeV
- Exactly one muon satisfying:

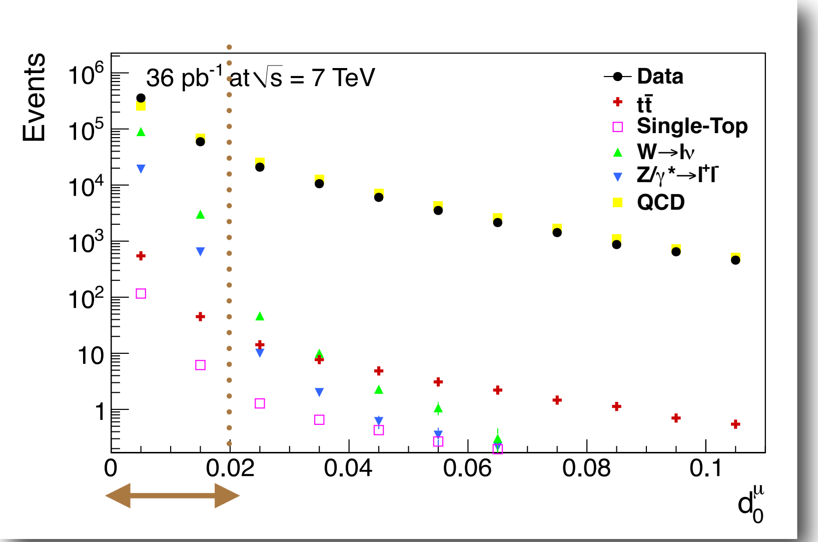
  - muon p_T > 20 GeV/c

  - located in the detector central region

  - muon ID requirements: Global and Tracker Muon, hits in the pixel and silicon tracker detectors, high quality tracks

  - the 2D impact parameter of the muon wrt the beam spot < 0.02 cm





Event Selection (II)

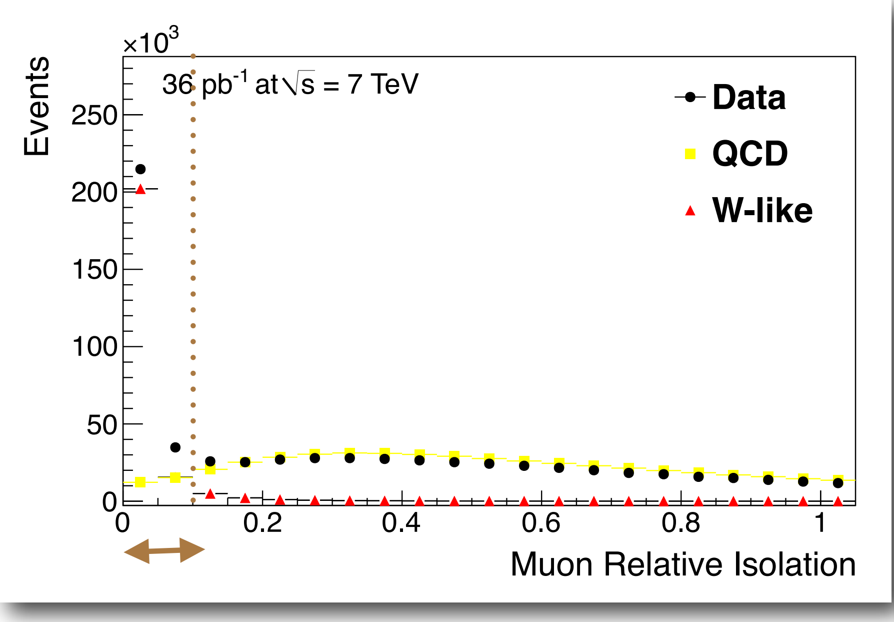
- The muon is isolated:

  - relative isolation: $$I^{\mu}_{rel} = \frac{I_{Trk} + I_{Ecal} + I_{Hcal}}{p_T^{\mu}} < 0.1$$

  where $$I_{Trk} = \sum_{\substack{i \neq j \\ \Delta R < 0.3}} (p_{T,i}^{track})$$

  $$I_{Ecal} = \sum_{\substack{i \\ \Delta R < 0.3}} (E_{T,i}^{Ecal}), \quad I_{Hcal} = \sum_{\substack{i \\ \Delta R < 0.3}} (E_{T,i}^{Hcal})$$

  - the muon is geometrically away from any energetic jet



Event Selection (III)

- Veto events with additional loose muons (looser selection)
- Veto events with isolated electrons
- At least 3 jets satisfying:

  - jet pT > 30 GeV/c

  - are located in the central detector region

  - jet ID (emf, at least 2 RecHits containing 90% of the jet energy, fraction of energy in the hottest HPD readout)

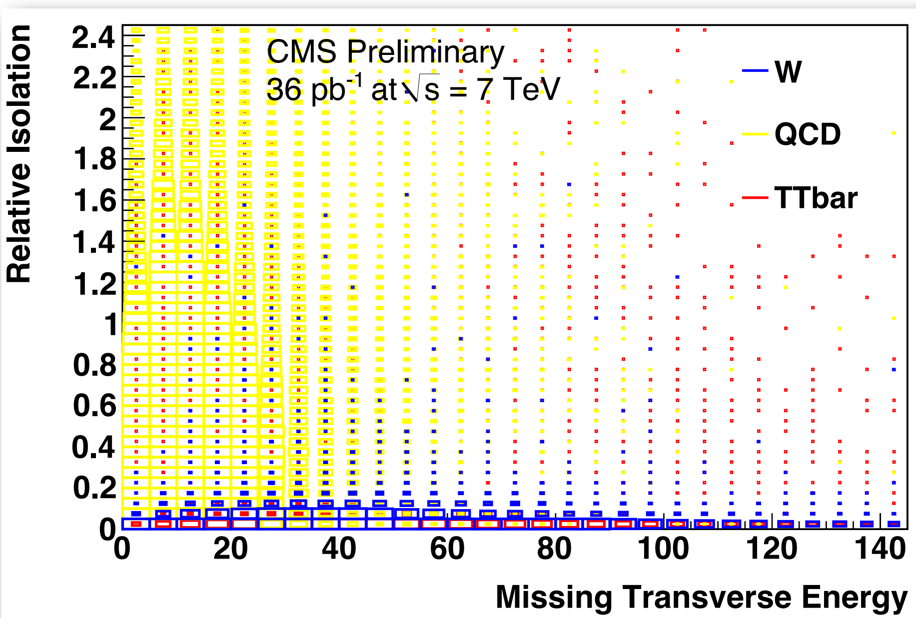
Cut	Number of events
Clean Filters	49034698
HLT	7727624
Good PV	7718212
One Isolated Muon	214368
Veto Loose Muon	208313
Veto Electron	207536
$\not{E}_T > 20$ GeV	157654
=1 jet	20012
=2 jet	4506
=3 jet	1111
$\geq$4 jets	459

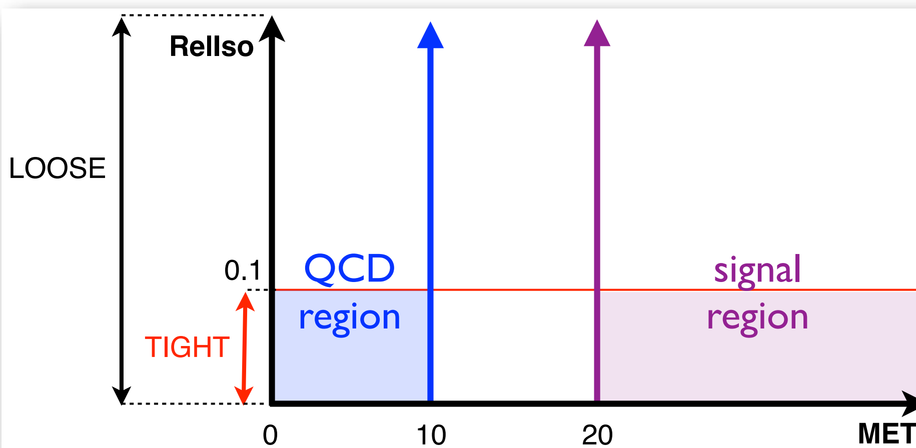
- keep 50% of signal (>= 3 jets)

~ 50% of the selected sample is background

QCD background estimation (I)

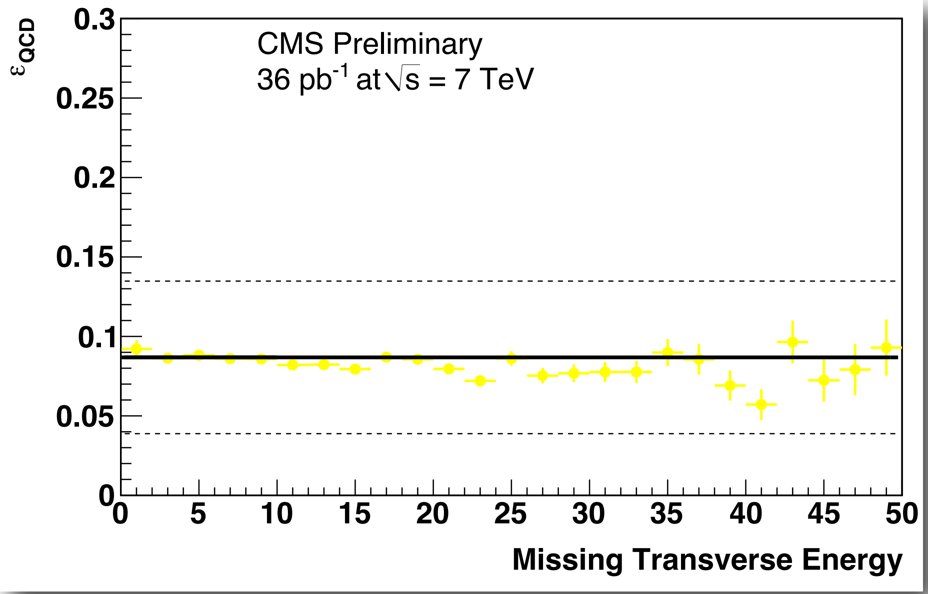
- The Matrix Method is used to estimate the QCD background before tagging
- The method uses the fact that muons from QCD events are not isolated
- Data-driven method: the key is to find a region with the minimum contamination from events with W bosons
- The method uses two samples: loose and tight. The difference between the samples is the muon relative isolation cut





QCD background estimation (II)

- We measure ϵ_{QCD} in the QCD dominated data sample as the efficiency of the muon from QCD sample to pass the isolation cut



- We measure ϵ_{sig} in the simulated sample (corrected to data)

- To correct the bias from the W-like contamination in the QCD region, the W(Z) contribution is subtracted from data in that region.

$$N_{Bkg}^{tight} = \epsilon_{Bkg} \frac{\epsilon_{sig}(N_1 + N_2) - N_2}{\epsilon_{sig} - \epsilon_{Bkg}}$$

where $N_1 = N_l - N_t$

$N_2 = N_t$

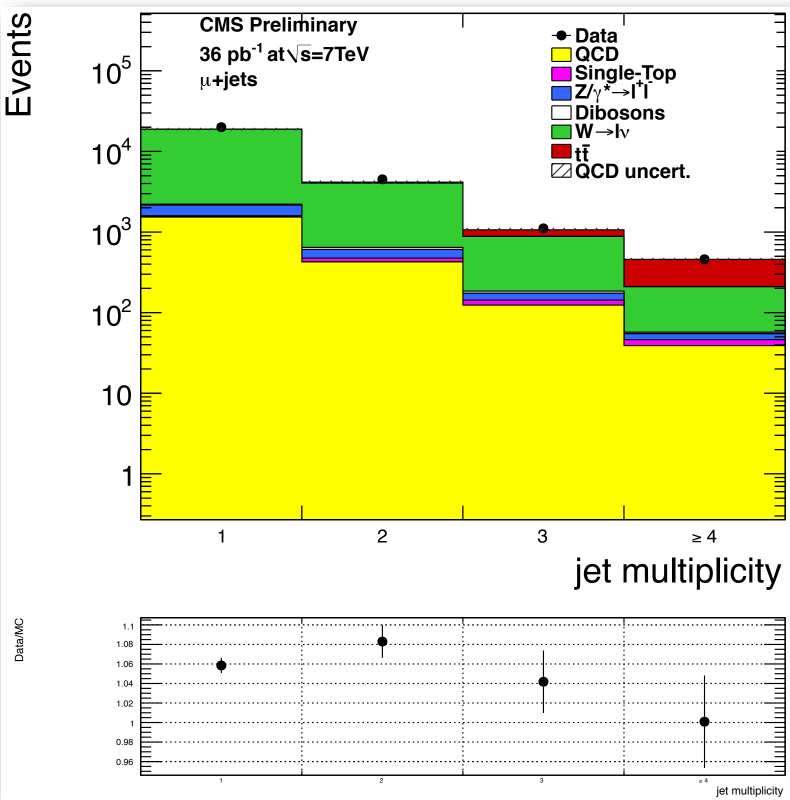
$$N_W^{tight} = \epsilon_{sig} \frac{N_2 - \epsilon_{Bkg}(N_1 + N_2)}{\epsilon_{sig} - \epsilon_{Bkg}}$$

$Njets$	default	W(Z) subtracted
$= 1$	0.231 ± 0.003	0.189 ± 0.002
≥ 2	0.209 ± 0.006	0.151 ± 0.006
$= 2$	0.208 ± 0.007	0.152 ± 0.006
$= 3$	0.211 ± 0.018	0.140 ± 0.016
≥ 4	0.242 ± 0.039	0.165 ± 0.035

Matrix Method results

- Matrix Method separates QCD from W-like.

Njets	N_{W-like}	N_{QCD}	ε_{sig}	ε_{QCD}
1	18358 ± 483	1654 ± 461	0.98 ± 0.02	$0.189 \pm 0.002 \pm 0.042$
2	4113 ± 192	393 ± 180	0.97 ± 0.02	$0.151 \pm 0.006 \pm 058$
3	1003 ± 60	108 ± 49	0.97 ± 0.02	$0.151 \pm 0.006 \pm 058$
≥ 4	426 ± 26	33 ± 15	0.96 ± 0.02	$0.151 \pm 0.006 \pm 058$



- The scale factors = ratio between the Matrix Method estimate and the MC simulation

Njets	SF_{W-like}	SF_{QCD}
1	0.92 ± 0.01	2.54 ± 0.12
2	0.97 ± 0.02	2.10 ± 0.19
3	1.08 ± 0.05	2.57 ± 0.47
≥ 4	0.96 ± 0.07	3.98 ± 1.55

- QCD, W and Z are scaled to the Matrix Method estimates

- Top pair and single top are normalized to the theoretical cross sections

W+jets background estimation

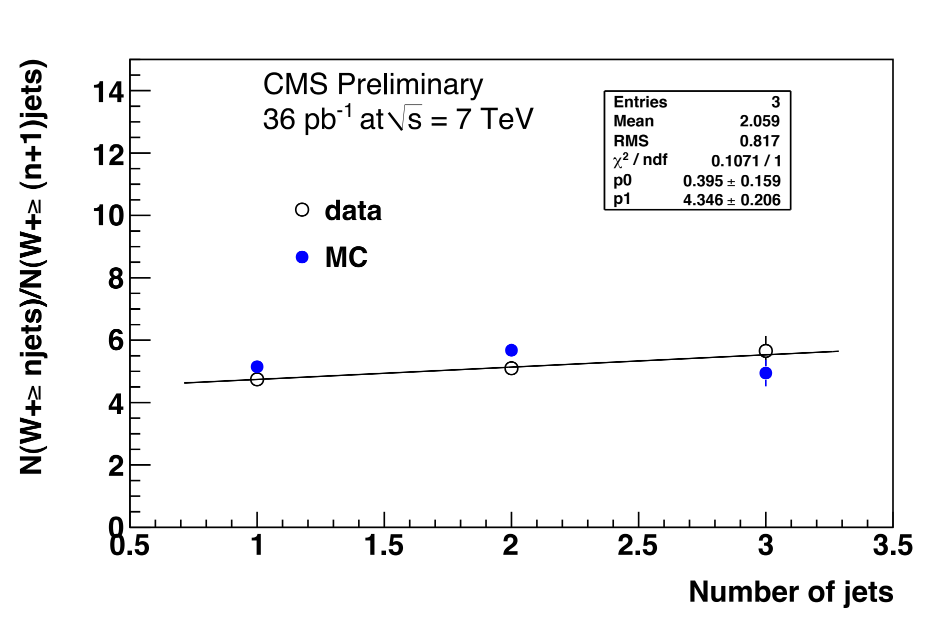
- The Berends Scaling method is used to estimate the W+jets background
- The ratio between the number of events with n and n+1 jets in W+jets should not depend on n

$$C(n) = \frac{N_W^{njets}}{N_W^{n+1jets}}$$

$$N_{W+jets}^{pretagged} = N_{data}^{pretagged} - N_{QCD,data-driven}^{pretagged} - N_{Z+jets,MC}^{pretagged} - N_{singleTop,MC}^{pretagged} - N_{WW,MC}^{pretagged} - N_{t\bar{t},MC}^{pretagged}$$

- We use C(n) = C(1) and use it to estimate the number of events in the 3rd and the 4th jet bin

$$N_{W+jets,\geq 3}^{pretagdata} = \frac{N_{W+jets,\geq 1}^{pretagdata}}{C(1)C(2)}$$



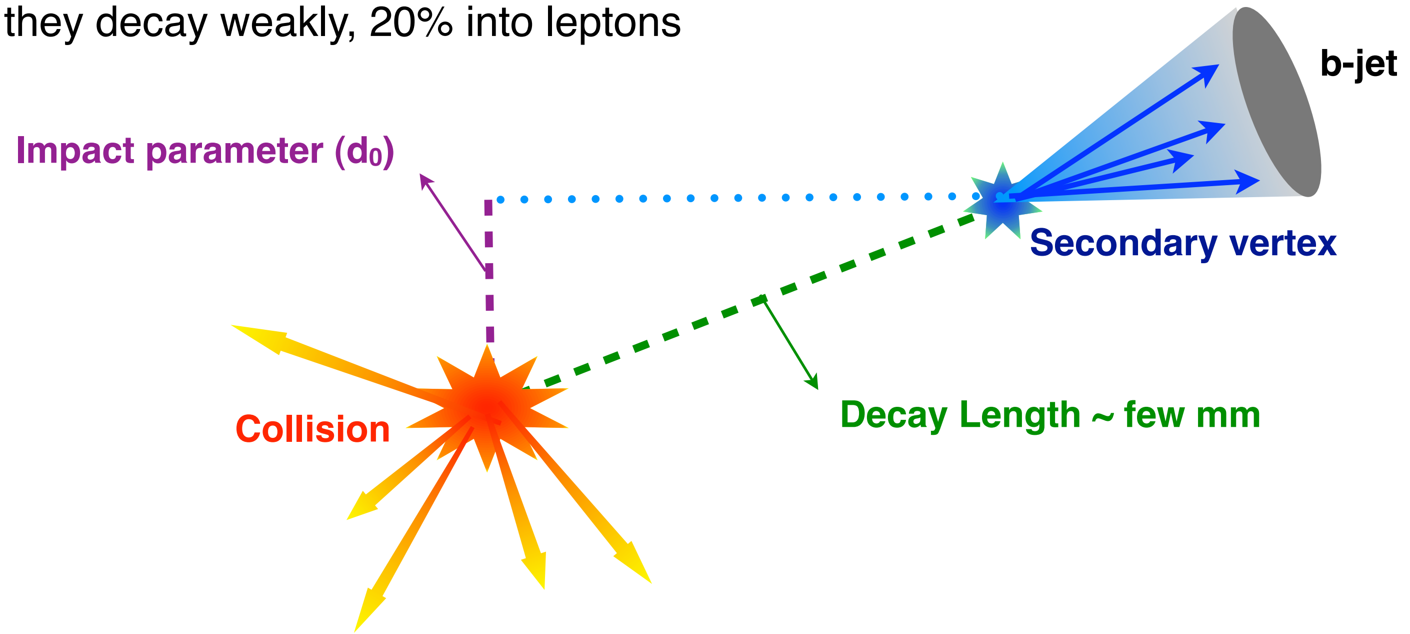
- After Berends scaling: 1 and 2 jet bins before tagging have Data/MC = 1 by construction

b-quarks Identification

- The tracks associated to the light quarks point to the interaction point (primary vertex).
- The b-quarks are “special”:

  - life time = 1.5 ps, they travel a few mm before decaying and a secondary vertex mass can be reconstructed

  - they decay weakly, 20% into leptons



- The *impact parameter* and the *flight length* can make the difference between the b-jets and the light quarks.

Simple Secondary Vertex Algorithm

- is based on the reconstruction of the secondary vertex
- discriminator - is defined using the 3D decay length

$$D = log(1 + \frac{\mid L_{3D} \mid}{\sigma_{L_{3D}}})$$

- N tracks are associated to the vertex

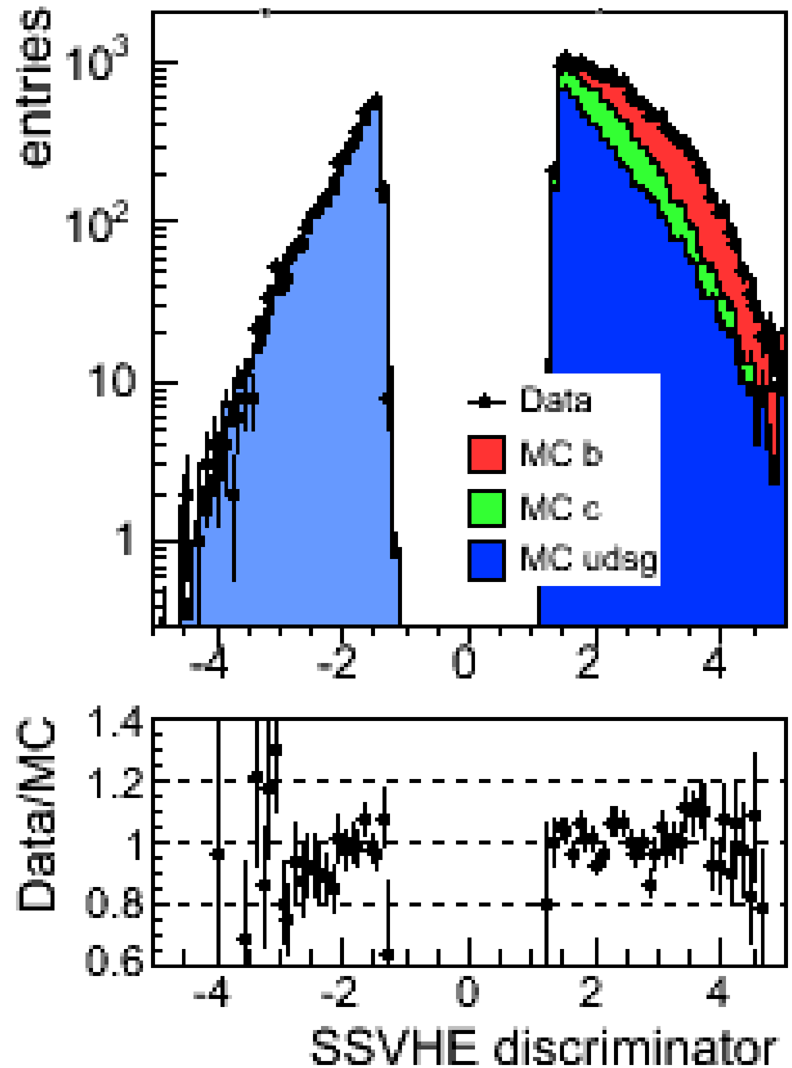
  - N >= 2, high efficiency

  - N >= 3, high purity

- operating points depending on light jets acceptance:

loose (L), medium(M), tight(T)

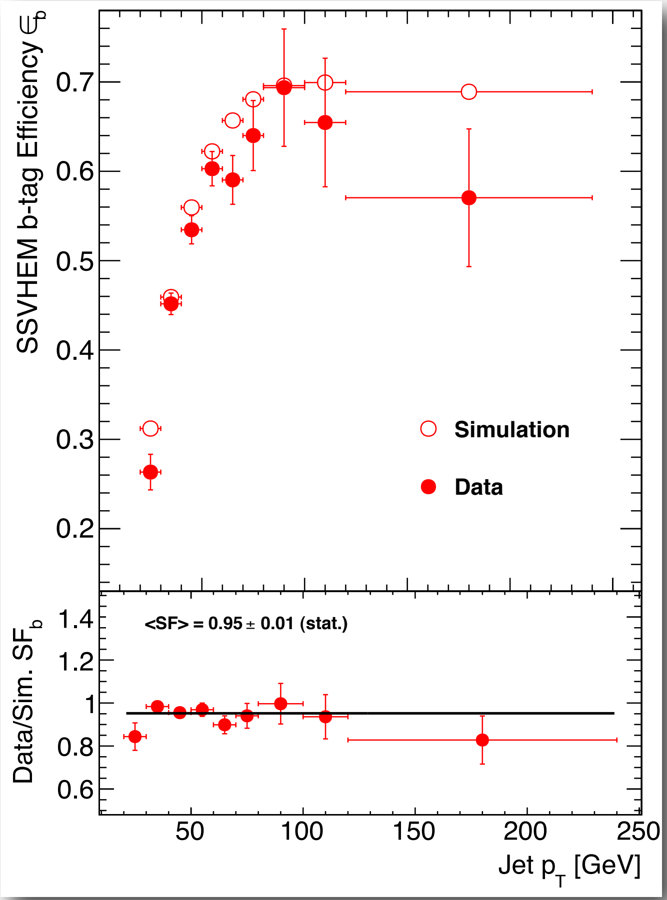
- we used: **SSVHE, medium operating point (1.74)**



Tagging jets

- Both data and Monte Carlo are tagged directly
- The tagging efficiency in MC is different from the tagging efficiency in data, for both heavy flavor and light jets
- MC needs to be corrected for that, using the scale factors
  - for tagged jets : $SF_x = \frac{\varepsilon_x^{Data}}{\varepsilon_x^{MC}}$
  - for non-tagged jets : $\bar{SF}_x = \frac{1-\varepsilon_x^{Data}}{1-\varepsilon_x^{MC}}$
- We used $SF_x = 0.9$ for b and c jets and parametrized values (vs transverse momentum and rapidity) for light jets
- The MC event is corrected by the event weight:

$$weight_{event} = \prod_{Ntaggedjets} SF_{taggedjet} * \prod_{Nnon-taggedjets} \bar{SF}_{non-taggedjet}$$



W+jets background

- The W+jets is the dominant background by far
- The W+jets sample contains jets with different flavors: b,c and light (u,d,s,g)
- The jet flavor separation is performed by matching the reconstructed jets with partons before hadronization with $\Delta R(jet, parton)$

  - **b jet** - if the jet is matched to a b parton, $\Delta R(jet, parton) < 0.5$

  - **c jet** - if the jet is matched to a c parton, $\Delta R(jet, parton) < 0.5$ and not a b parton

  - **light jet** - if the jet is matched to neither a b or c parton
- Flavor History Filter built in CMS

N jets	Wb(b)	Wc(c)	W+light
≥ 1	0.008	0.160	0.806
≥ 2	0.016	0.166	0.776
≥ 3	0.025	0.162	0.746
≥ 4	0.031	0.146	0.726
$= 1$	0.007	0.159	0.813
$= 2$	0.014	0.167	0.783
$= 3$	0.024	0.166	0.751

Fraction of events with jets with different flavor composition for the W+jets sample

Heavy flavor scale factor

- After Berends scaling, the pre-tagged samples have the ratio Data/MC = 1 in the 1st and 2nd jet multiplicity bins, by construction
- After applying the b-tagging, the normalization is not preserved
- We attribute this discrepancy to the predicted rate of Wb(b) and Wc(c) events

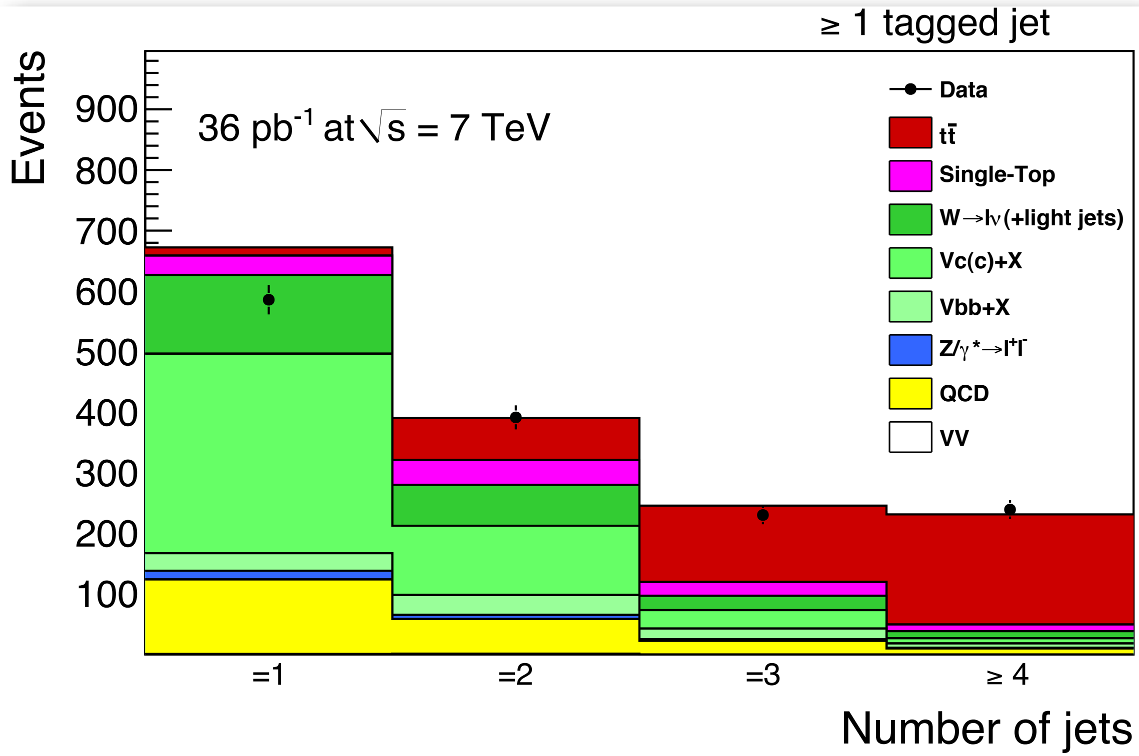
$$W + jets = k * Wb(b) + k * Wc(c) + Wlight$$

- We calculate the heavy flavor scale factor in the =2 jets multiplicity bin, =1 tag and we found that we need k ~ 2.0 to preserve the normalization Data/MC
- We used the same k = 2+/- 1 for >= 3 jets multiplicity bin.

Sample	k = 1	k = 2.0
Data	**360**	**360**
QCD	55.27	55.27
W+b(b)	14.29	28.57
W+c(c)	56.73	113.46
W+light	66.78	66.78
Z+jets	6.63	6.63
WW	2.37	2.37
single Top	37.18	37.18
Total bkg	239.25	310.26
TTbar	55.48	55.48
Total MC	**294.73**	**365.74**
Data/MC	**1.22**	**0.98**

Results after tagging

- After tagging, the signal dominates in the 3 and 4 jet multiplicity bins



Sample	≥ 3 jets, ≥ 1 tag
Data	471
QCD	32.1
$Wb\bar{b}$ + jets	24.7
$Wc\bar{c}$ + jets	38.9
W+light	35.1
Z+jets	3.9
WW	1.3
single top	33.8
Total background	169.7
$t\bar{t}$	307.6
Total MC	477.3
Data/MC	1.01

* only statistical errors are shown

- QCD - is normalized to the Matrix Method results obtained before tagging
- W+jets - obtained from Berends scaling
- Top - normalized to the theoretical predictions $\sigma_{t\bar{t}} = 157^{+23}_{-24}$ pb, http://arxiv.org/abs/hep-ph/1007.3492

with m_t = 172.5 GeV. Uncertainties are obtained with PDFs: MSTW2008, CTEQ6.6 and NNPDF2.0

- We used k = 2.0 as the heavy flavor correction for Wb(b) and Wc(c) after tagging

Top pair cross section extraction

- We extract the cross section in the >= 3 jets and >= 1 tag as the excess of data over background prediction

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N_{bkg}}{BR \times L \times \epsilon^{pres} \times P^{tag}}$$

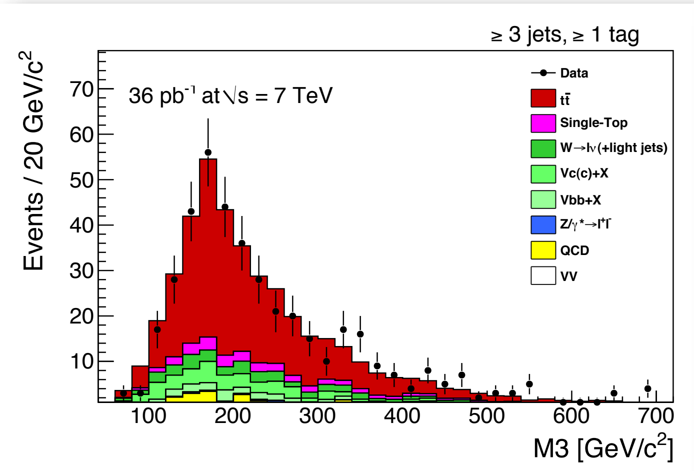
- Efficiency of the pre-tagged selection (measured in MC simulated $t\bar{t}$ sample): $BR \times \epsilon_{sel} = 0.076$
- Tagging probability (>= 1 tag): $P_{\geq 1\ \mathrm{tag}} = 0.718$

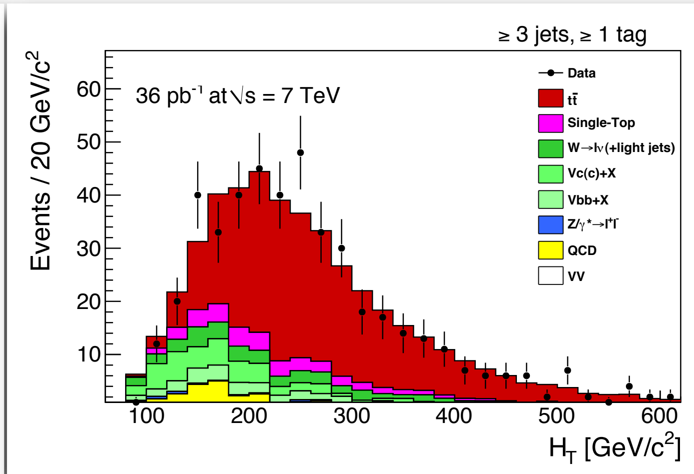
$$\mathcal{L} = 36.15\ \mathrm{pb}^{-1}$$

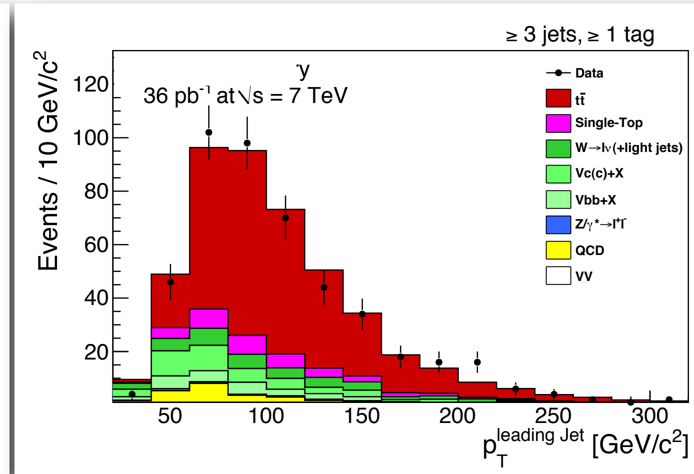
Systematic	$\sigma_{t\bar{t}}$		Fractional changes (%)	
QCD estimation (50% variation)	+ 0.9	-1.2	+ 0.5	- 0.8
theoretical cross section (30% variation for TTbar) (50% variation for single Top)	+ 4.3	- 4.4	+ 2.8	- 2.9
W+jets estimation(Berends scaling)	+ 0.0	- 8.3	+ 0.0	-5.4
b-tagging Efficiency (15% variation for $SF_{b/c}$) (DataBase variation for SF_{light}) *very conservative (all up, all down)	+ 18.7	- 13.2	+ 12.2	- 8.6
HF scale factor (50% variation)	+ 15.2	-15.6	+ 9.9	- 10.2
JES	+ 7.8	- 8.2	+ 5.1	-5.3
Q^2 normalization (W+jets, Z+jets, TTbar)	+ 9.9	- 3.64	+ 6.4	- 2.4

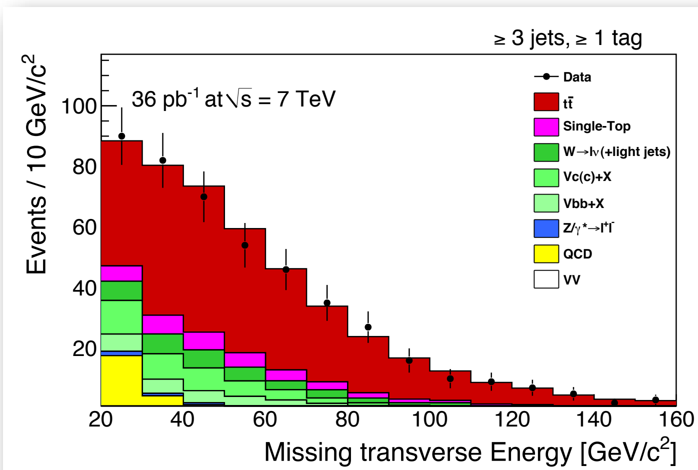
$$\boxed{\sigma_{t\bar{t}} = 155 \pm 11(\mathrm{stat.})^{+28}_{-24}(\mathrm{syst.}) \pm 6(\mathrm{lumi.})\mathrm{pb}}$$

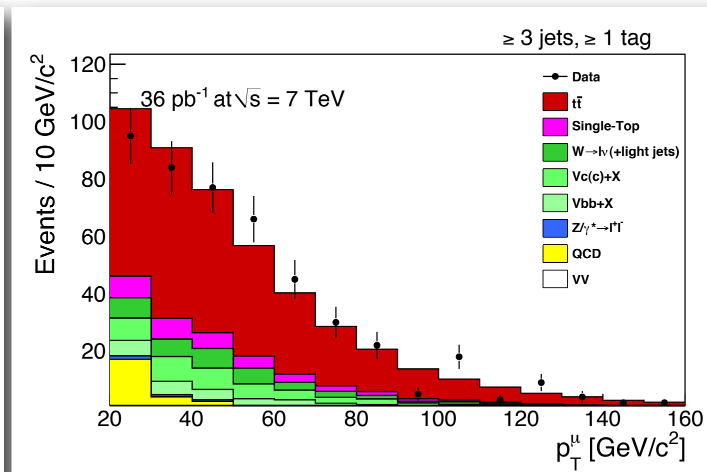
Sanity Checks









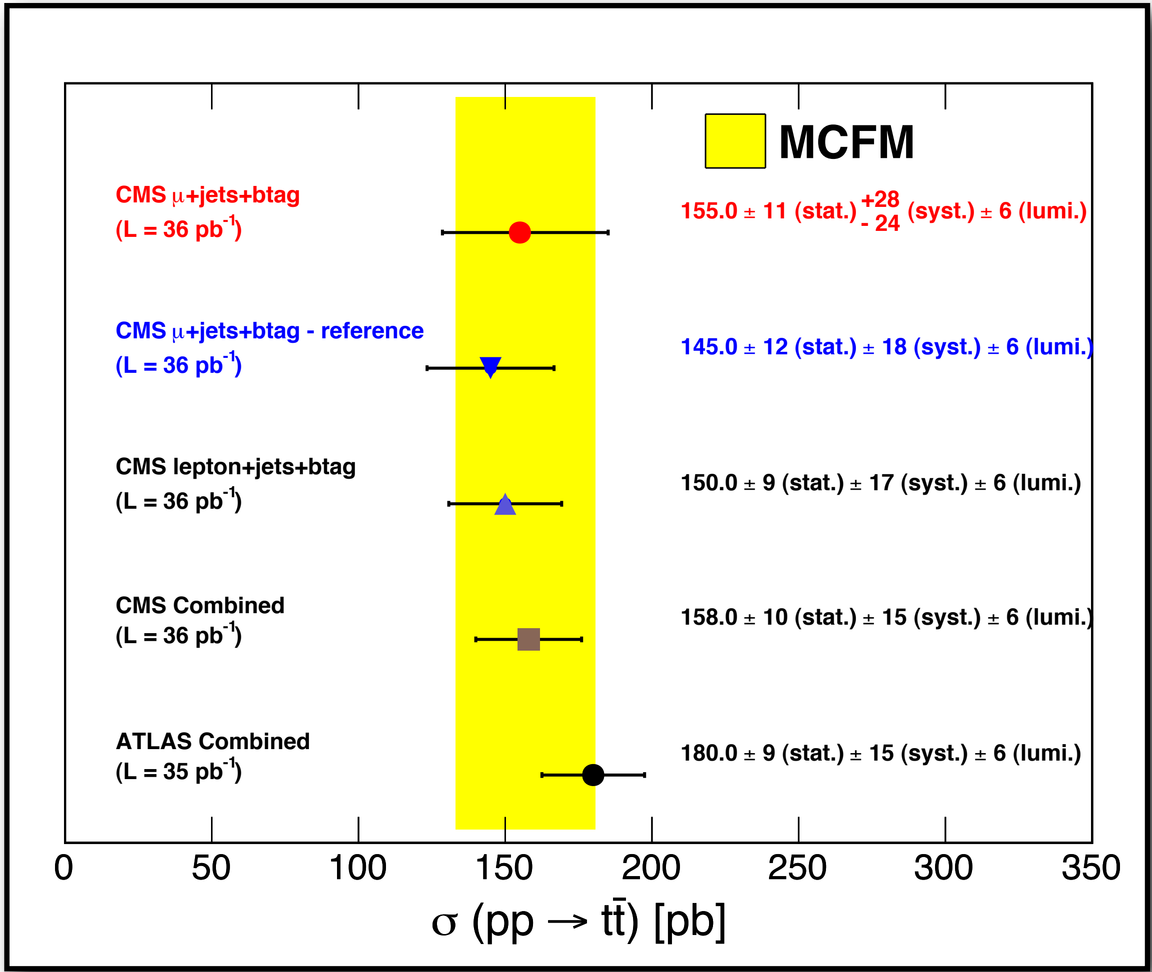


* only statistical errors are shown

- Top - normalized to the measured cross section value

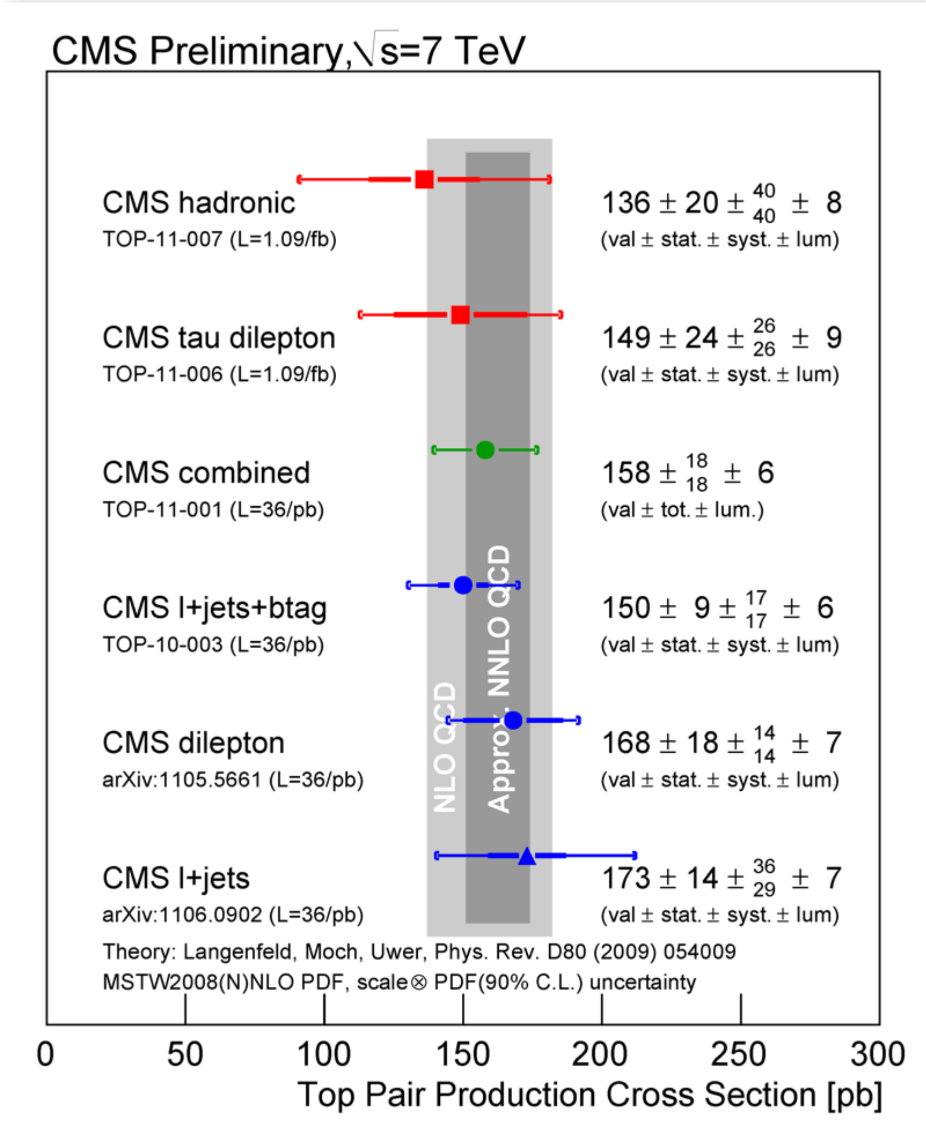
Conclusions

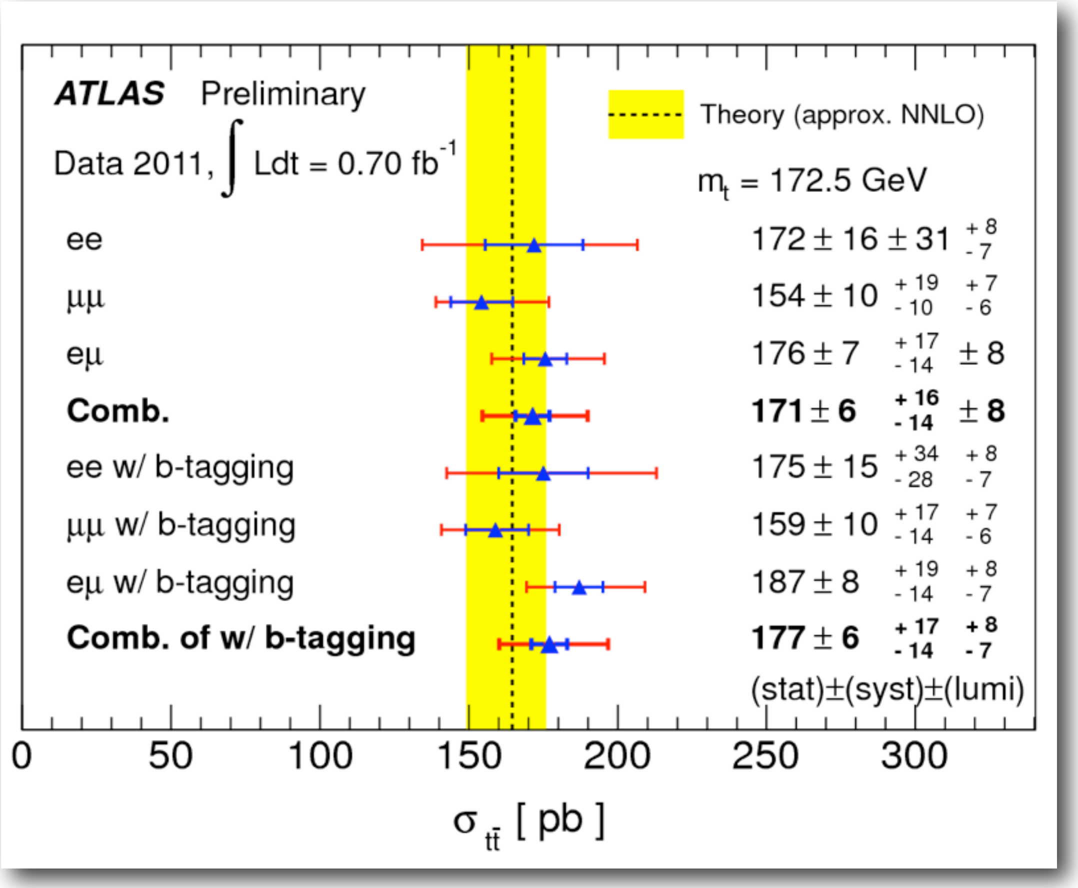
- A counting measurement of the Top pair production cross section has been shown.
- We used data-driven techniques to estimate the amount of QCD and W+jets before tagging
- A heavy flavor scale factor was measured for Wb(b) and Wc(c) in the 2 jets multiplicity bin, with =1 tag. This correction was applied for 3 and 4 jet bins.



- This measurement is quoted as a cross check for the reference analysis in ***CMS TOP PAS-10-003***
- The result is in agreement with the results obtained with different techniques and different channels.
- The measured top quark pair production cross section at CMS agrees within uncertainties with the result obtained at ATLAS and with the theoretical prediction.

Latest Public Results

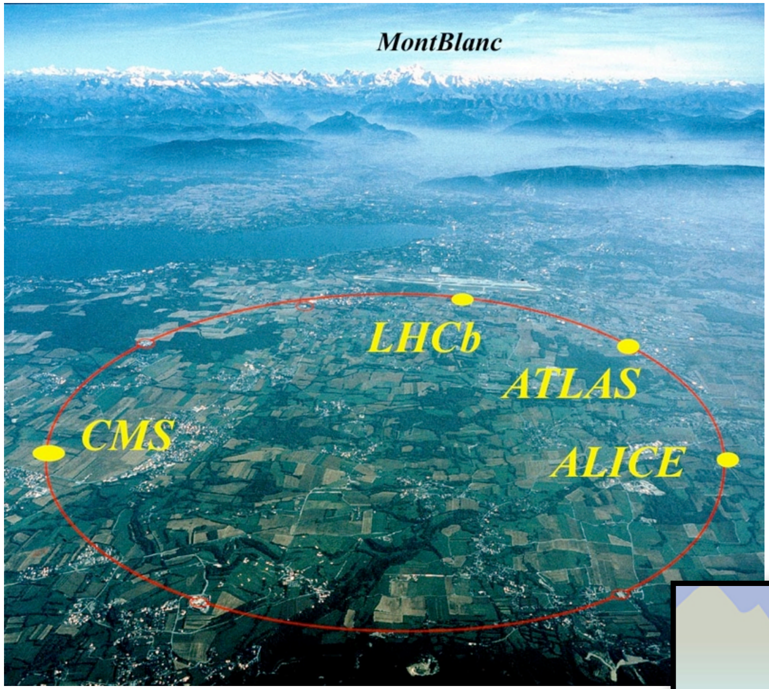




Back up

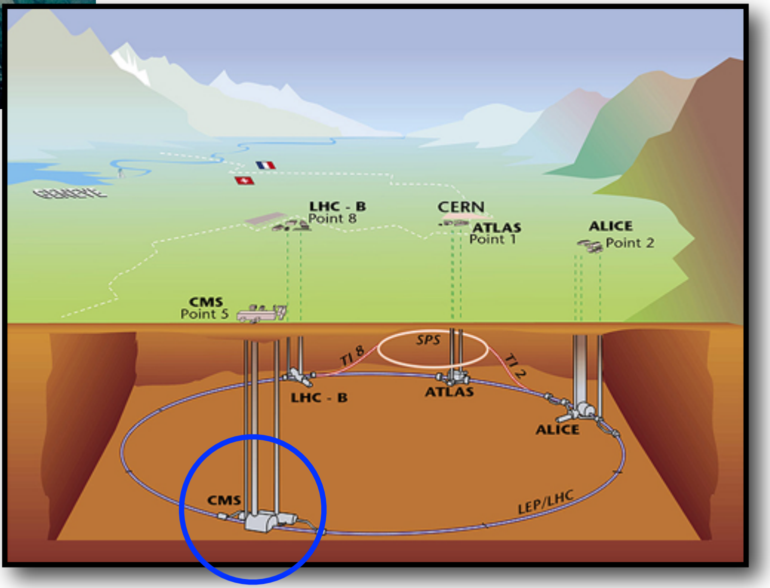
LHC

★ LHC (Large Hadron Collider) : the biggest hadron collider in the world



- located on the border of France and Switzerland
- 27 km circumference

- designed to collide protons at 14 TeV center of mass energy

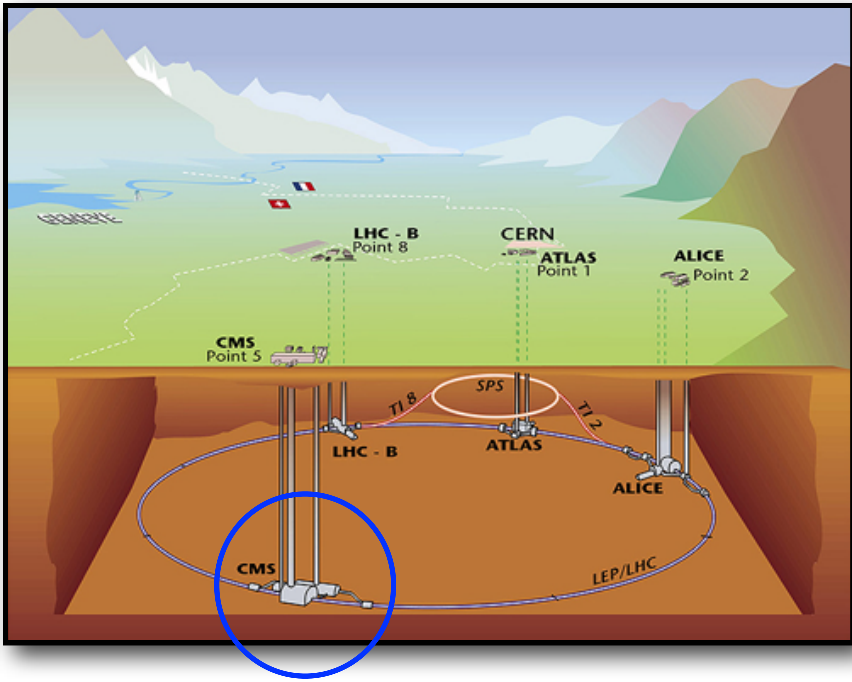


CMS

- Large Hadron Collider (LHC) - run at 7 TeV center of mass energy since 2010

LHC and CMS

LHC

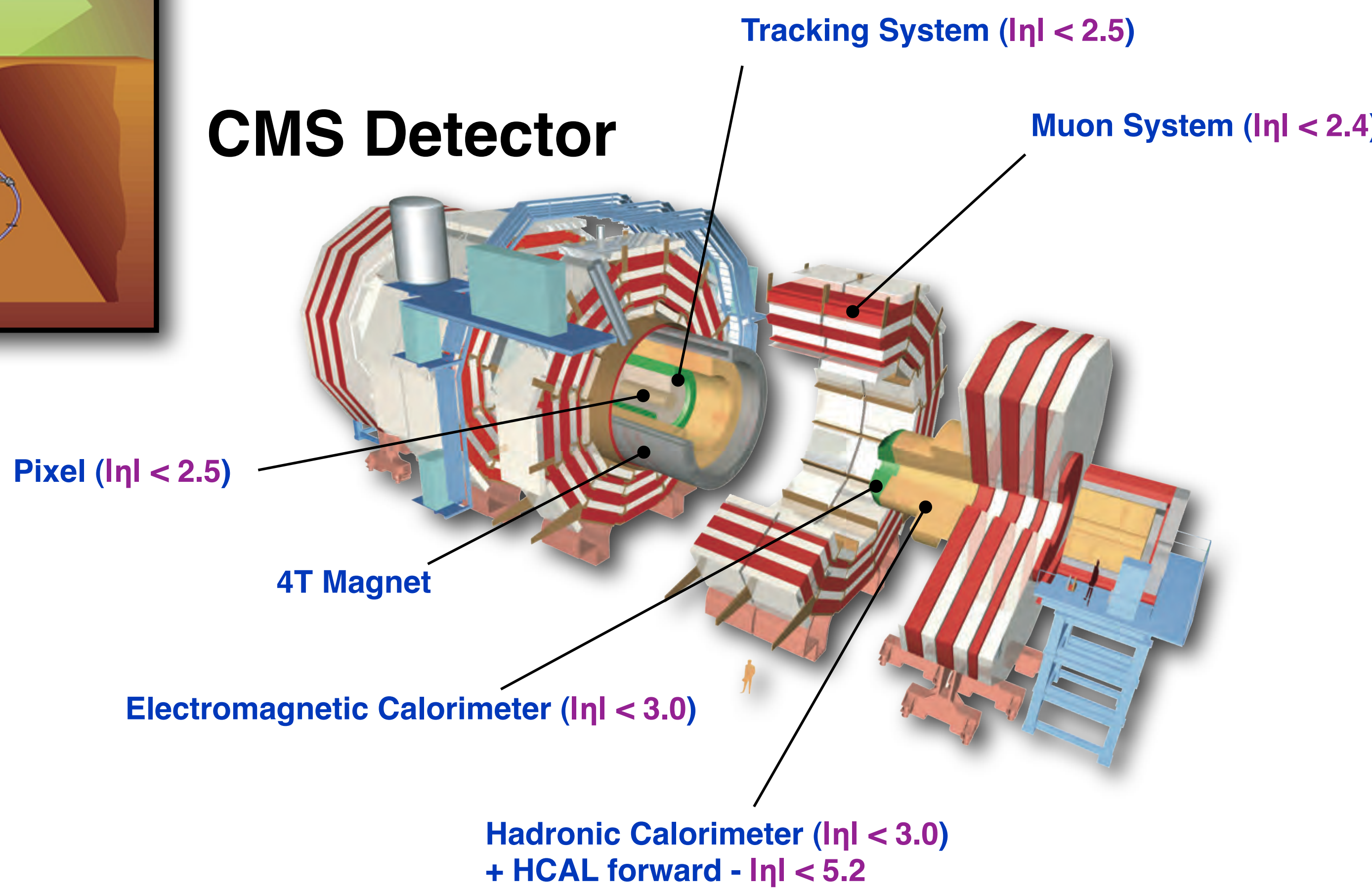


- CMS (Compact Muon Solenoid) characteristics:

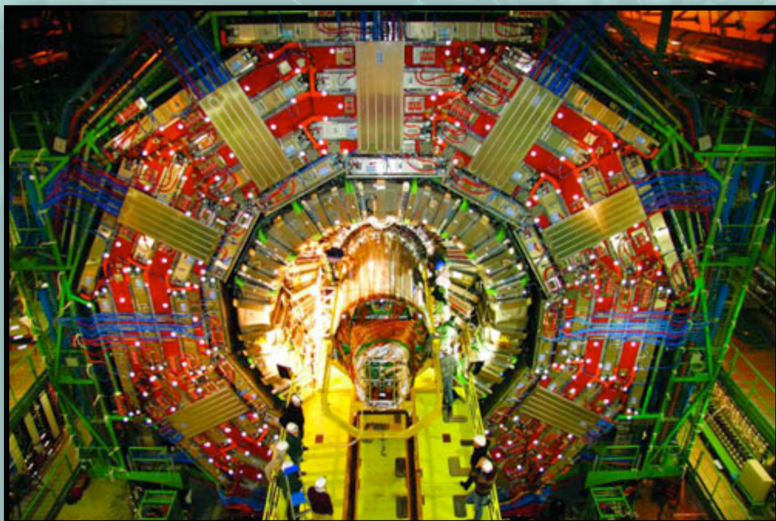
  - CMS Tracker has the largest silicon area ever built, providing a very fine granularity - helping in identifying the b-quarks

  - excellent instrument in muon identification

CMS Detector



CMS Detector



Solenoid

- very good muon identification
  - $|\eta| < 2.4$

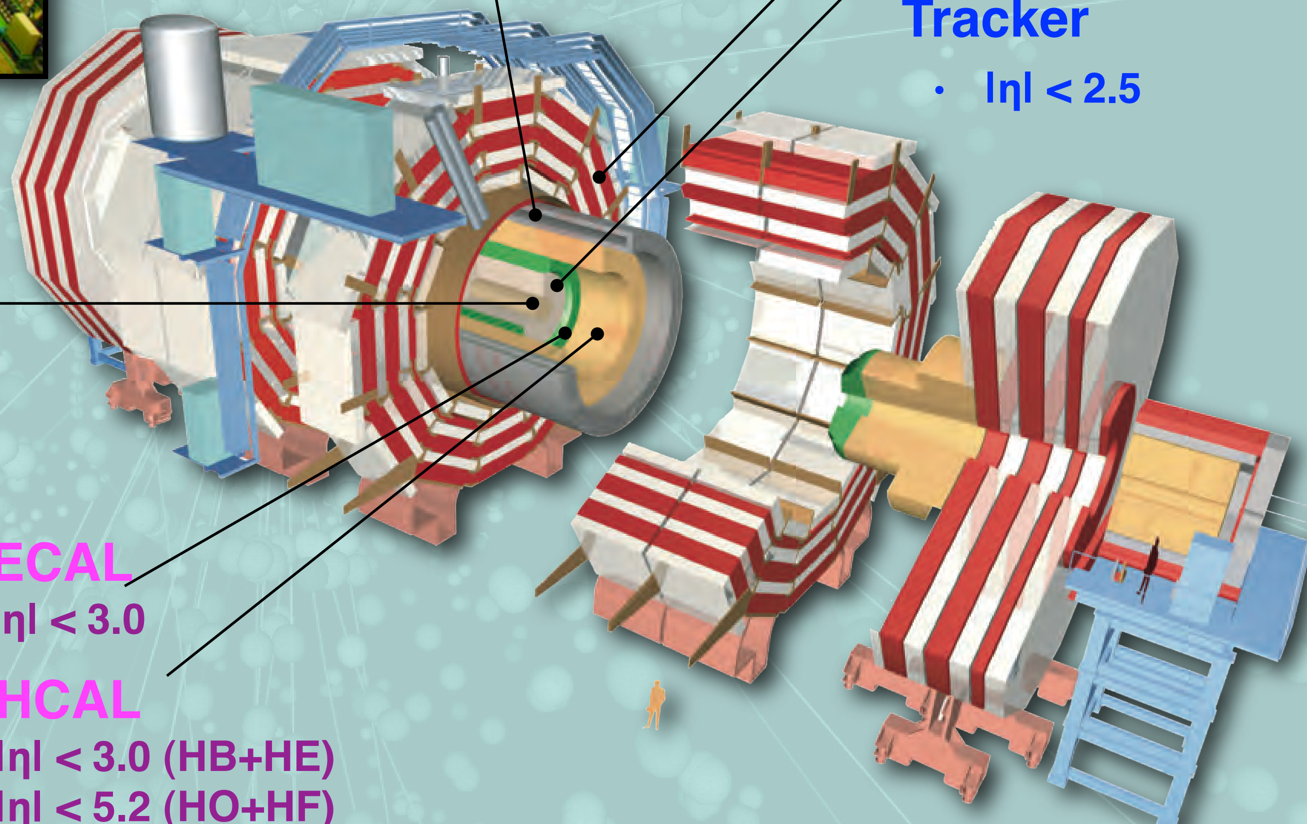
Muon System

Silicon Strip Tracker

- $|\eta| < 2.5$

3D vertex reconstruction
- very important for identification of the b-jets
  - $|\eta| < 2.5$

Pixel



- good measurement of hadron jets and MET

ECAL
$|\eta| < 3.0$

HCAL
$|\eta| < 3.0$ (HB+HE)
$|\eta| < 5.2$ (HO+HF)

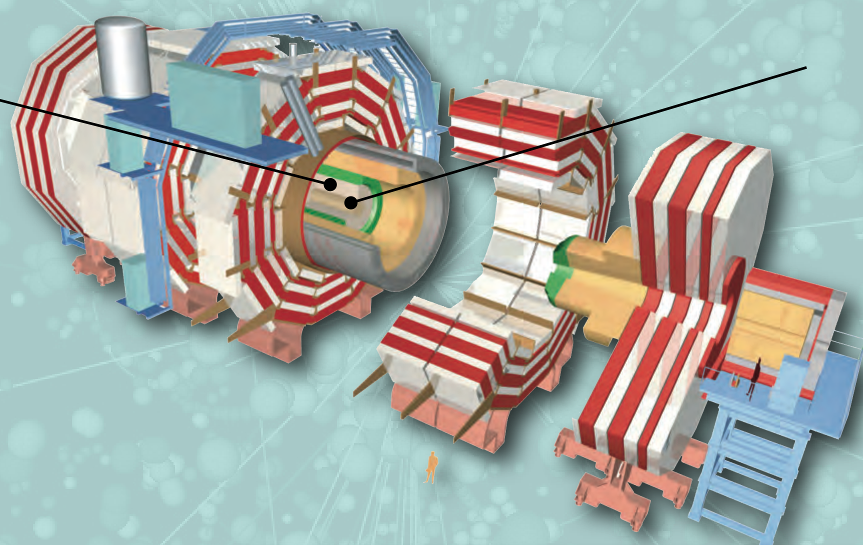
CMS Subdetectors (1)

★ CMS Tracking System requirements:

- high granularity : it must measure precisely and efficiently the charged particles tracks and also provide a precise reconstruction of the secondary vertices
- fast response
- radiation hardness

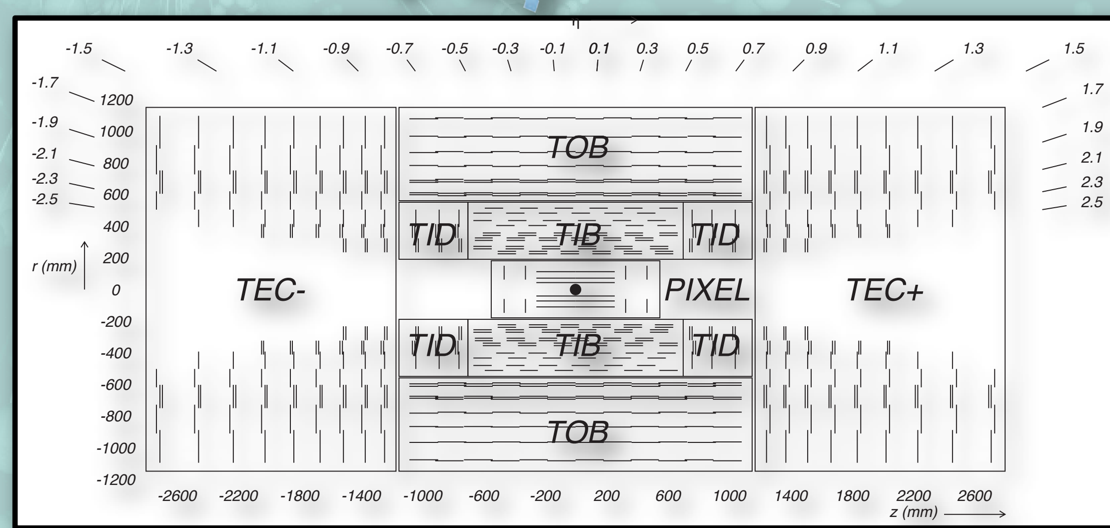
Silicon Tracker

- 198 m^2 active silicon area
- 10 layers of silicon microstrip detectors in the barrel: **TIB** (4) & **TOB** (6)
- 11 disks in the endcaps: **TID** (3) & **TEC** (9)
- $|\eta| < 2.5$
- S/N > 10 is expected for 10 years of operation at the designed CM energy
- some of the layers are double in order to provide a measurement in a z-direction (barrels) and r-direction (disks)
- provides a resolution at the order of hundreds of microns



Pixel

- 3 layers in the barrel
- 2 disks in the endcaps
- **$|\eta| < 2.5$**
- very good track resolution on all 3 directions: a 3D vertex reconstruction is possible - very important for SV reconstruction from b and tau decays



CMS Subdetectors (2)

Electromagnetic Calorimeter (ECAL)

- Uses lead tungstate ($PbWO_4$) crystals
- ECAL barrel (**EB**): **$|\eta| < 1.48$**
- ECAL endcaps (**EE**): **$1.48 < |\eta| < 3.0$**
- **Preshower** - to improve position resolution: **$1.65 < |\eta| < 2.6$**

ECAL Design Purpose:

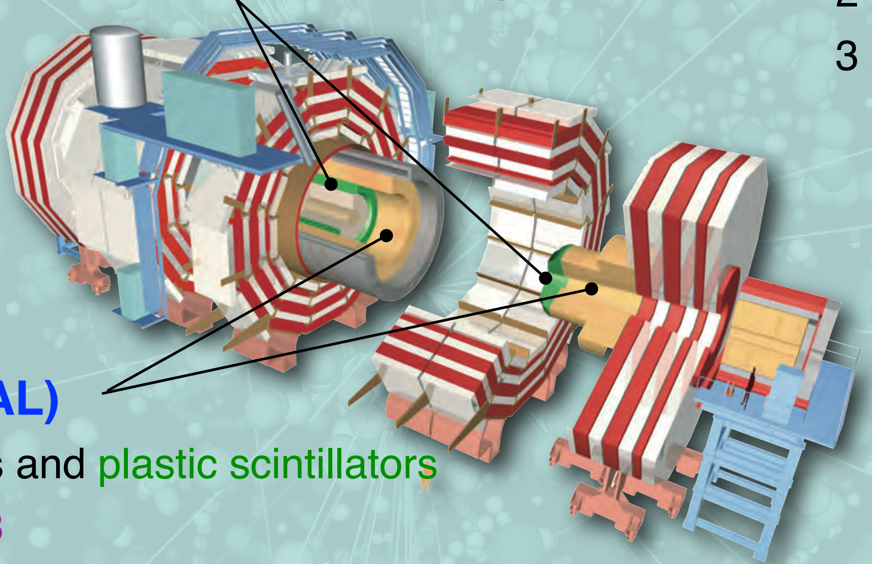
- detect H decay into 2 photons

$X_0 = 0.89$ cm

25.8 X_0 (barrel)

24.7 X_0 (endcap)

3 X_0 (preshower)



Hadronic Calorimeter (HCAL)

- Uses steel/brass absorbers and plastic scintillators
- HCAL barrel (**HB**): **$|\eta| < 1.3$**
- HCAL endcaps (**HE**): **$1.3 < |\eta| < 3.0$**
- HCAL outer barrel (**HO**) - outside of the solenoid: **$|\eta| < 1.3$**
- HCAL forward (**HF**) - steel absorbers and quartz fibers: **$3.0 < |\eta| < 5.2$**

HCAL Design Purpose

- measures the hadron jets and neutrinos

$\lambda_1 = 16.42$ cm

5.39 λ_1 - 10.3 λ_1 (barrel)

10 λ_1 (endcap)10 λ_1 (forward)

CASTOR

- **$5.2 < |\eta| < 6.6$**

ZDC

- **$|\eta| > 8.3$**

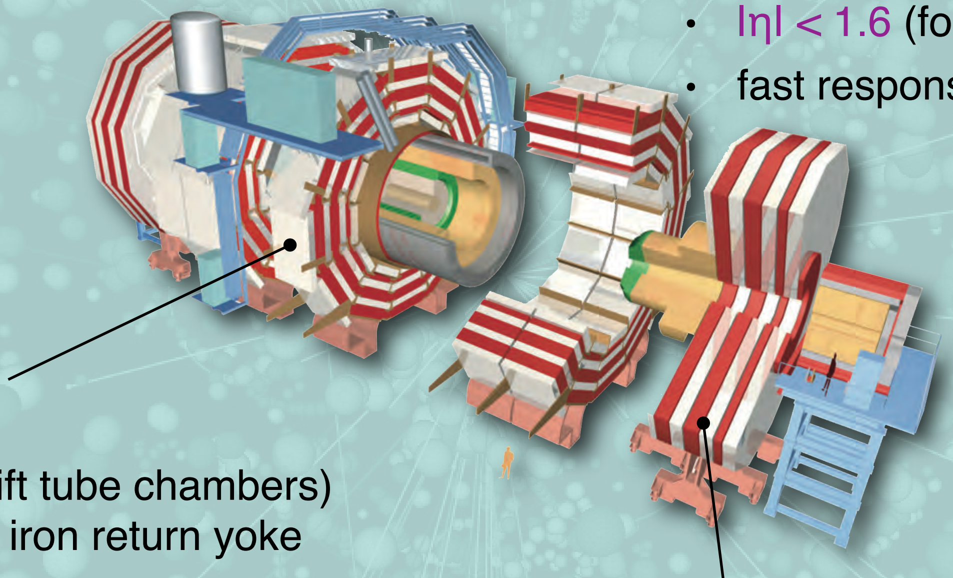
CMS Subdetectors (3)

Muon System Functions

- muon identification and measurement
- triggering
- rejection of the non-muon background

RPC (2 end-caps)

- resistive plate chambers embedded into the iron yoke layers
- trigger system added in both DT and CSC
- $|\eta| < 1.6$ (for now)
- fast response, good time resolution



DT (barrel)

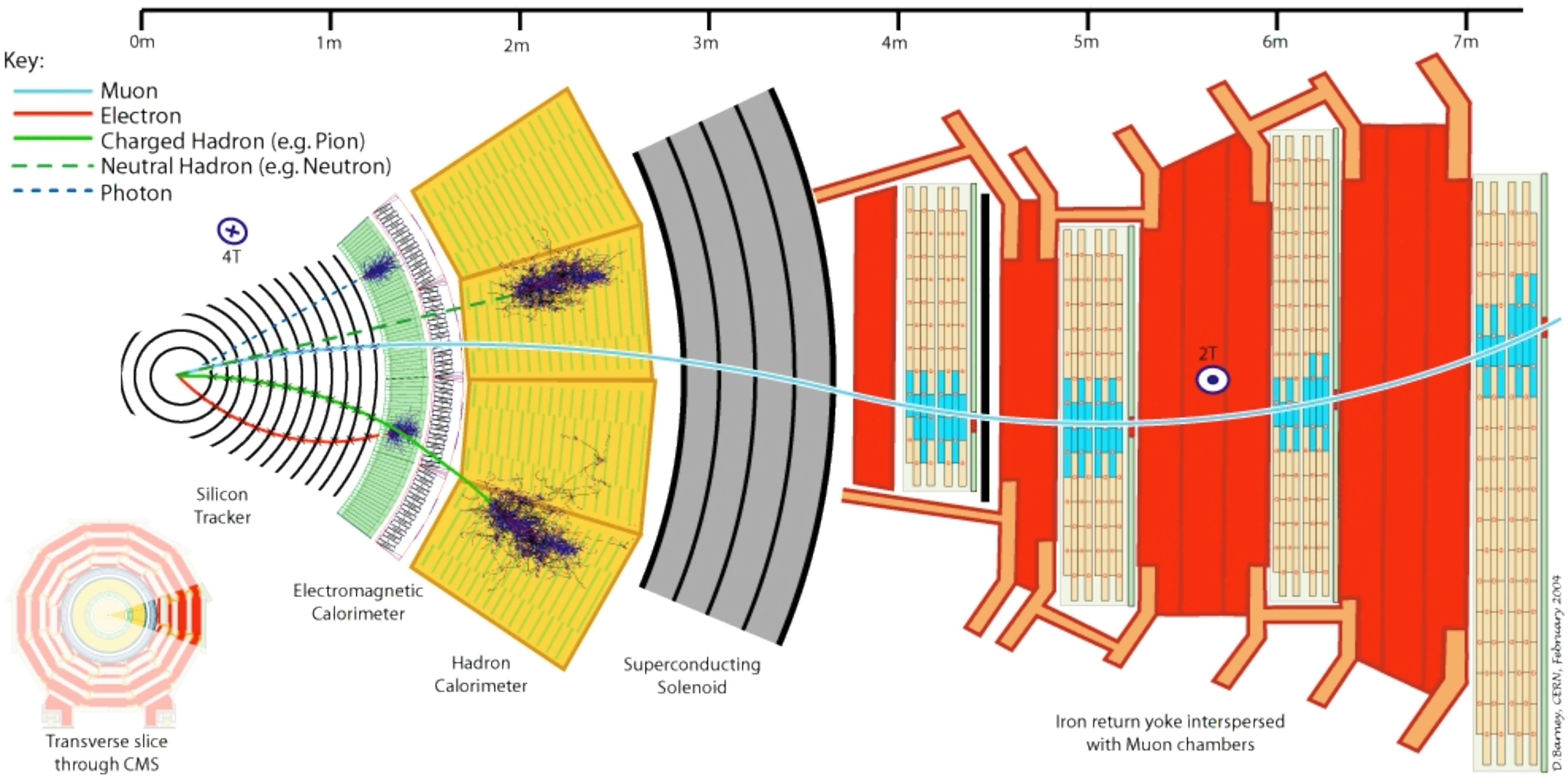
- 4 muon stations (drift tube chambers) interleaved with the iron return yoke plates
- $|\eta| < 1.2$
- good efficiency for linking together muon hits from different stations into a single muon track

CSC (2 end-caps)

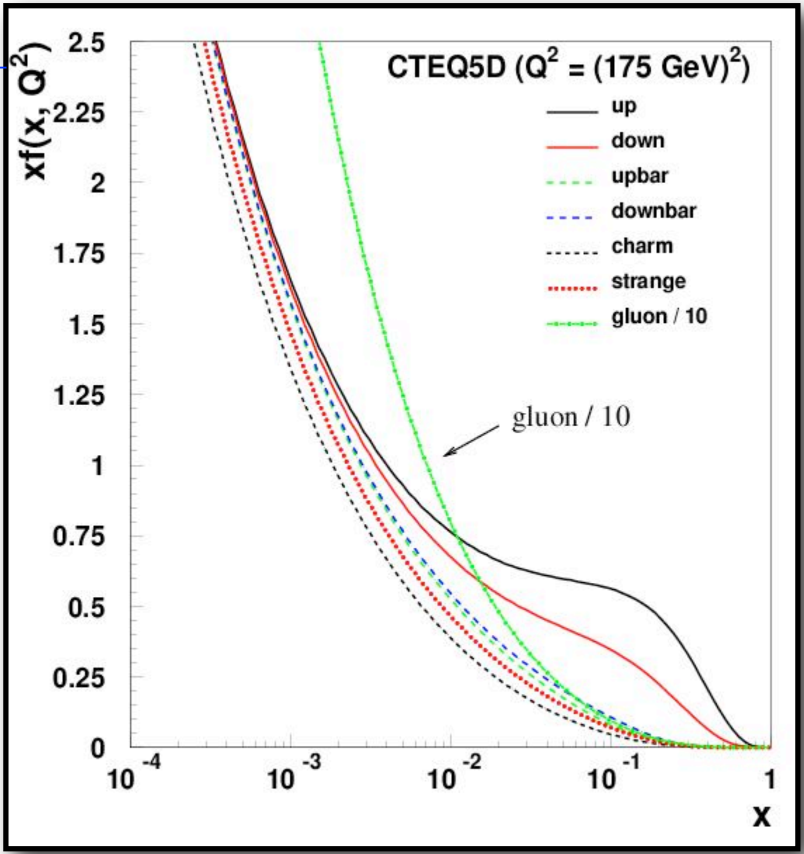
- 8 muon stations (cathode strip chambers) interleaved with the iron return yoke plates
- $0.9 < |\eta| < 2.4$
- efficient matching of hits to those in other stations and to the CMS inner tracker

How does it work ?

Tracks reconstruction, energy deposits and vertices identification
are the keys for particles identification

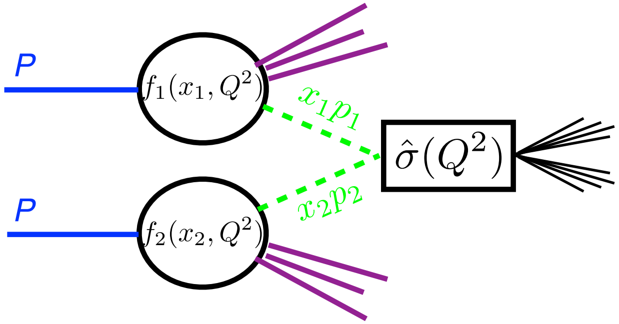


$m_t = 175.5 \pm 4.6(stat) \pm 4.6(syst)$ *CMS* http://arxiv.org/pdf/1105.5661

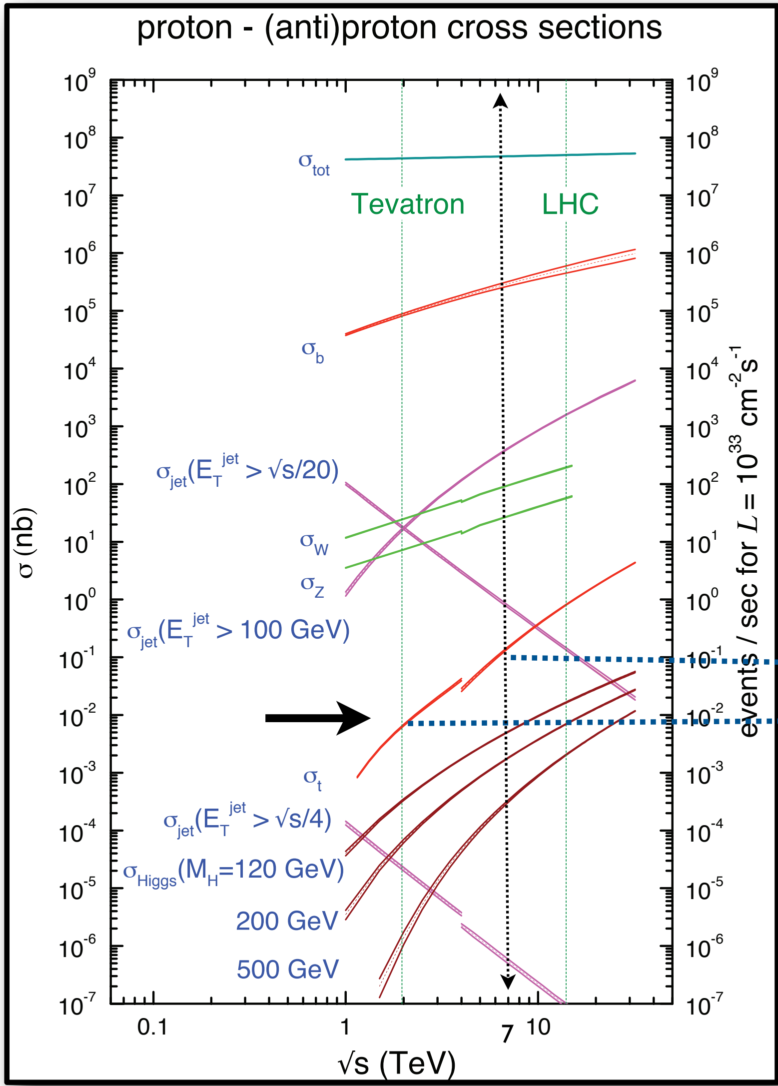


$$x \approx \frac{2m_t}{\sqrt{s}}$$

$$x \approx 0.05 \text{ at LHC } \sqrt{(s)} = 7 \text{ TeV}$$



How Much Top Quark at LHC ?

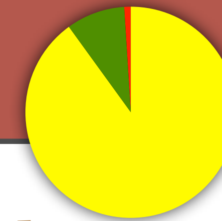


For 7 TeV :
20 x Tevatron XS

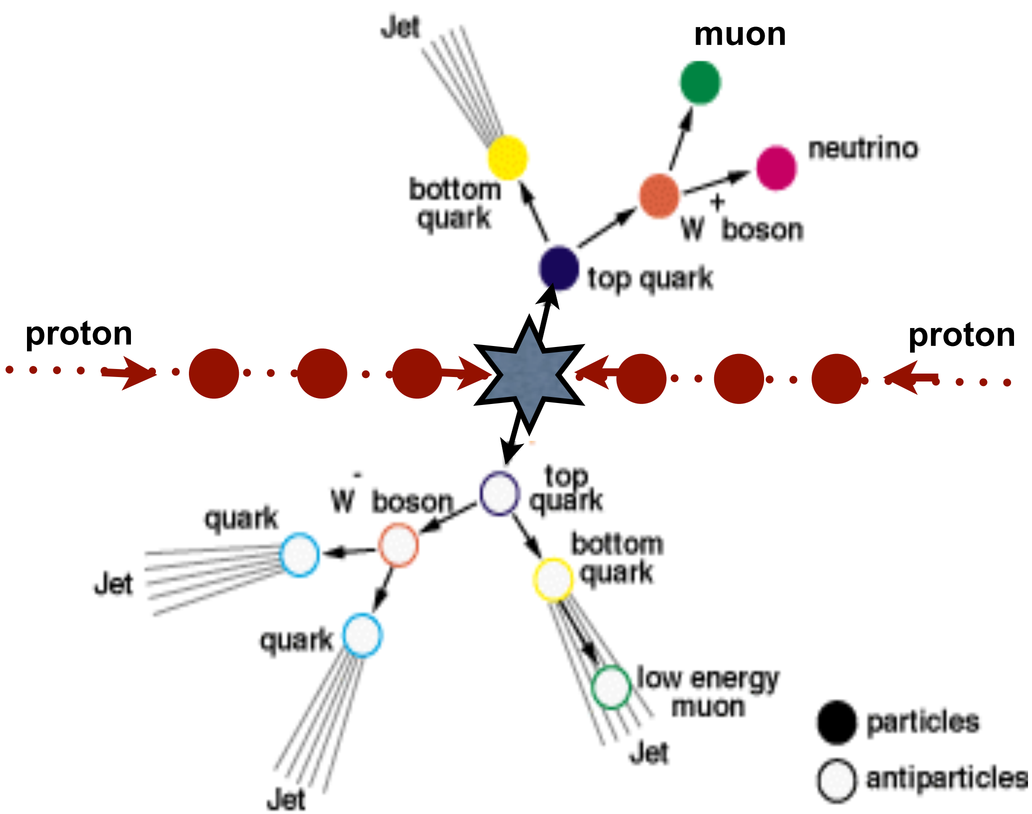
LHC is a Top factory

Event Selection

Data: 36 pb^{-1}



Muon+jets channel

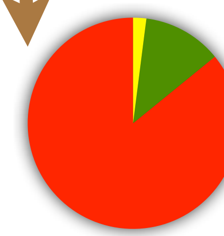


- **Event selection**:

  - *particle flow jets** reconstructed using anti-k_T algorithm (3 or more jets)
  - exactly one muon originating from the PV

  - muon isolation + missing transverse energy (neutrino)

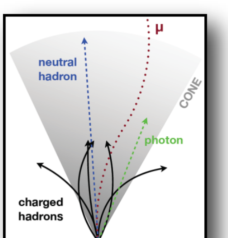
  - at least one jet is a b-jet



QCD

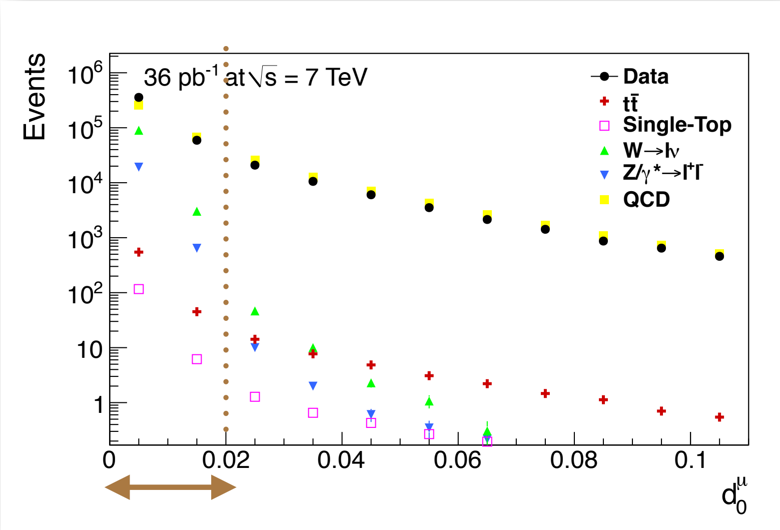
W-like

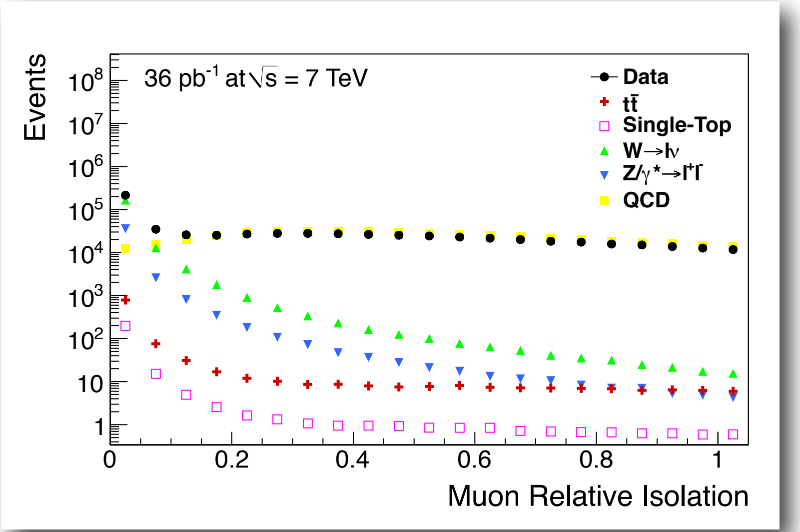
Top pair



***Particle flow jets** - from identified particles using all detector components

Event Selection

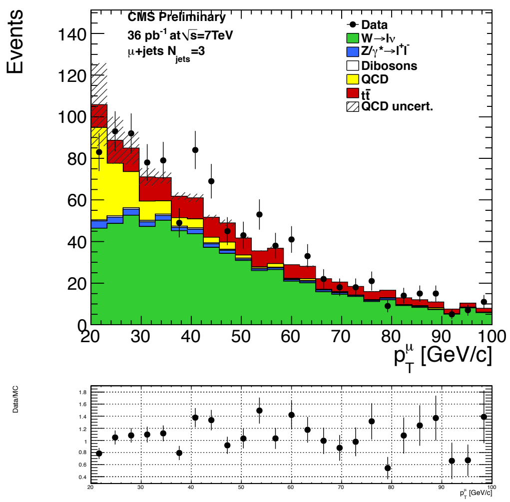


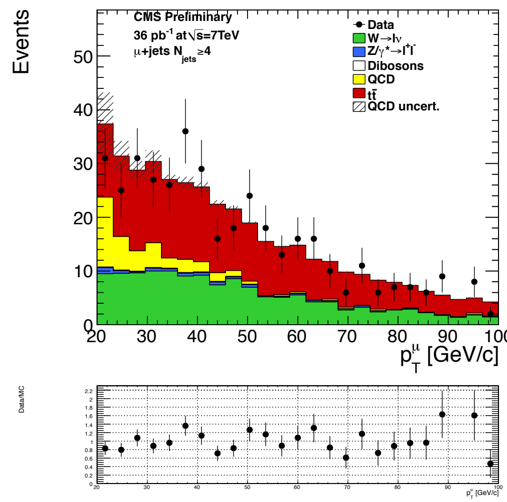


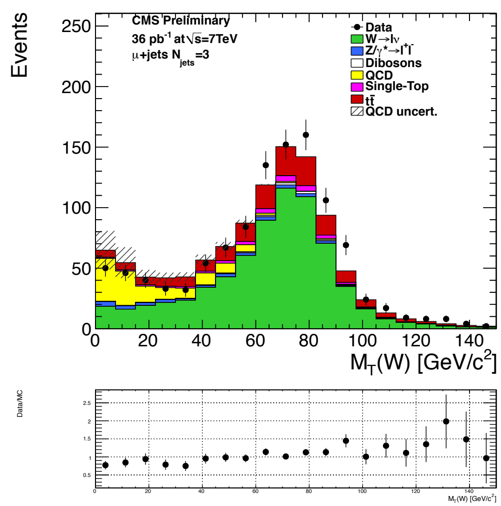
Event Selection - MC results

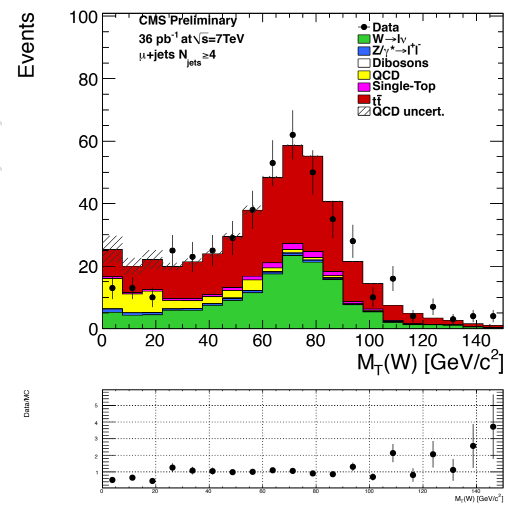
Cut	$t\bar{t}$	$W \to l\nu$	$Z \to l^+l^-$	QCD	t channel	tW channel	s channel	WW	Total
Processed	5670.0 ± 5.0	1127304.0 ± 293.0	109728.0 ± 68.8	3048454.8 ± 561.2	753.5 ± 1.1	381.6 ± 0.5	54.0 ± 0.1	1548.0 ± 1.1	4293893.9 ± 636.9
HLT	1784.8 ± 2.8	253269.6 ± 138.9	34376.0 ± 38.5	2586537.3 ± 517.0	261.7 ± 0.6	104.3 ± 0.3	19.7 ± 0.0	272.2 ± 0.5	2876625.7 ± 536.7
Good PV	1784.4 ± 2.8	252315.1 ± 138.6	34301.6 ± 38.5	2585475.4 ± 516.8	261.5 ± 0.6	104.3 ± 0.3	19.7 ± 0.0	271.7 ± 0.5	2874533.6 ± 536.5
One Iso mu	765.1 ± 1.8	174836.8 ± 115.4	14210.1 ± 24.8	19928.9 ± 45.4	157.3 ± 0.5	54.0 ± 0.2	10.7 ± 0.0	184.3 ± 0.4	210147.2 ± 126.5
Loose mu veto	741.7 ± 1.8	174829.4 ± 115.4	9283.8 ± 20.0	19906.5 ± 45.4	156.9 ± 0.5	52.4 ± 0.2	10.7 ± 0.0	179.4 ± 0.4	205160.8 ± 125.6
Electron veto	643.0 ± 1.7	174440.2 ± 115.2	9063.8 ± 19.8	19735.6 ± 45.2	153.8 ± 0.5	44.9 ± 0.2	10.5 ± 0.0	158.8 ± 0.3	204250.6 ± 125.4
met$>$ 20 GeV	578.4 ± 1.6	157490.0 ± 109.5	5968.4 ± 16.0	1121.5 ± 10.8	136.8 ± 0.5	39.4 ± 0.2	9.3 ± 0.0	137.1 ± 0.3	165480.8 ± 111.3
1 jet	32.6 ± 0.4	19209.8 ± 38.2	665.6 ± 5.4	647.8 ± 8.2	51.8 ± 0.3	6.3 ± 0.1	2.9 ± 0.0	51.7 ± 0.2	20668.5 ± 39.5
2 jets	116.2 ± 0.7	3900.6 ± 17.2	148.5 ± 2.5	185.9 ± 4.4	47.8 ± 0.3	13.1 ± 0.1	4.2 ± 0.0	36.9 ± 0.2	4453.1 ± 18.1
3 jets	184.6 ± 0.9	680.8 ± 7.2	31.8 ± 1.2	41.5 ± 2.1	18.6 ± 0.2	12.2 ± 0.1	1.3 ± 0.0	9.9 ± 0.1	980.7 ± 7.7
$\geq$4 jets	241.9 ± 1.0	180.2 ± 3.7	9.8 ± 0.7	8.3 ± 0.9	7.2 ± 0.1	6.8 ± 0.1	0.3 ± 0.0	2.4 ± 0.0	456.9 ± 4.1

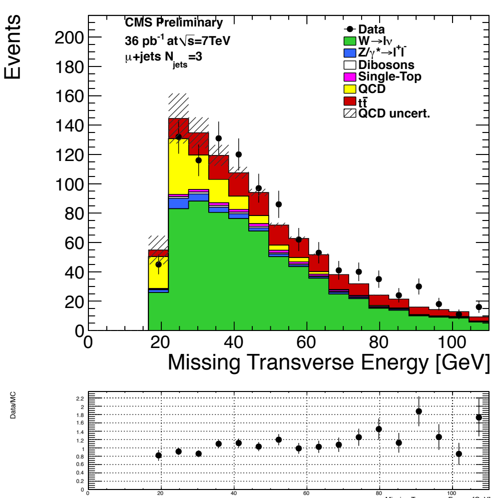
Matrix Method results - pre-tagged

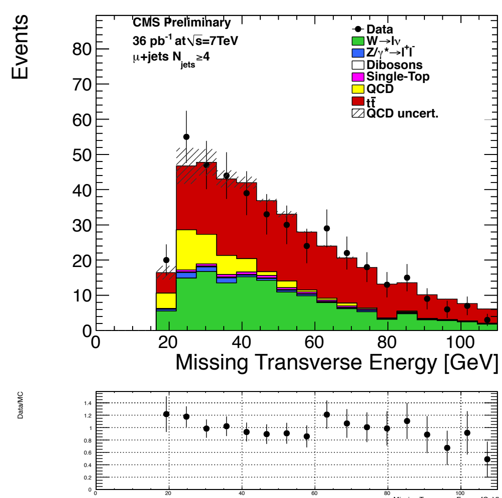




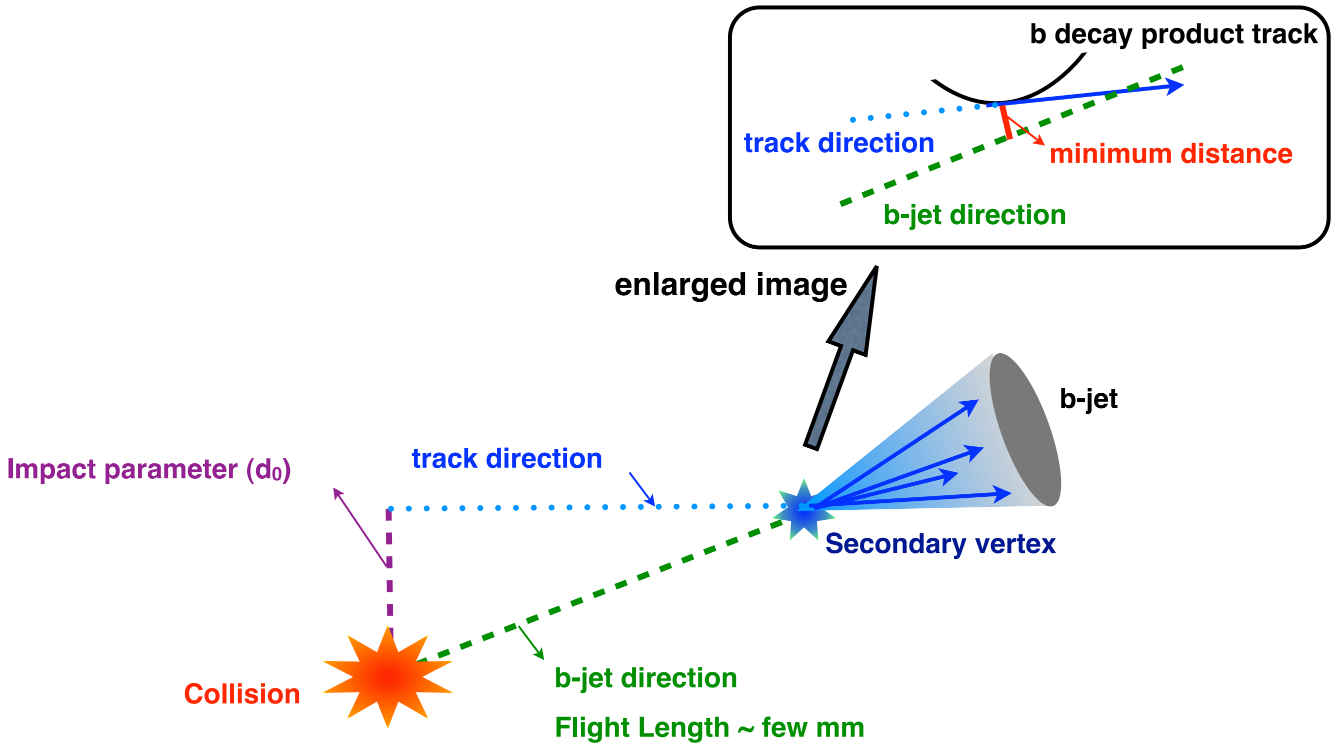








Impact Parameter

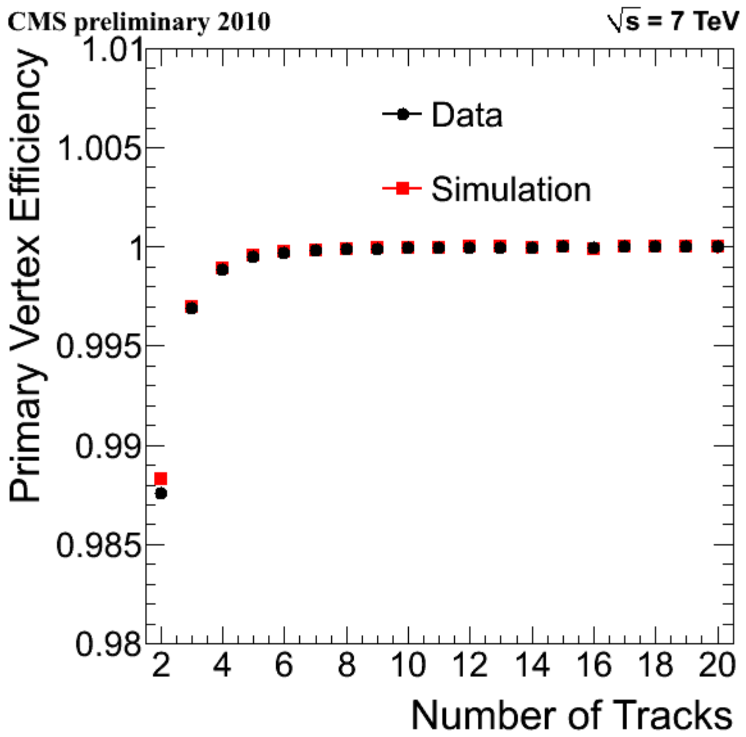


Vertex Reconstruction

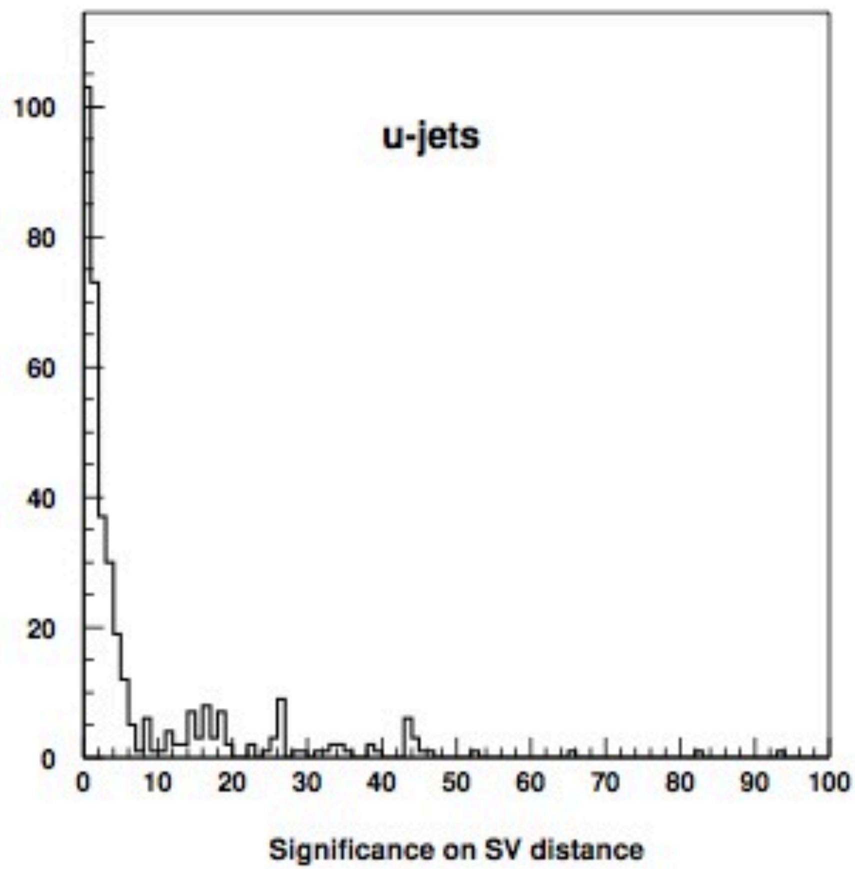
- efficiency and resolution for vertex reconstruction is strongly dependent on the number of tracks associated to the vertex

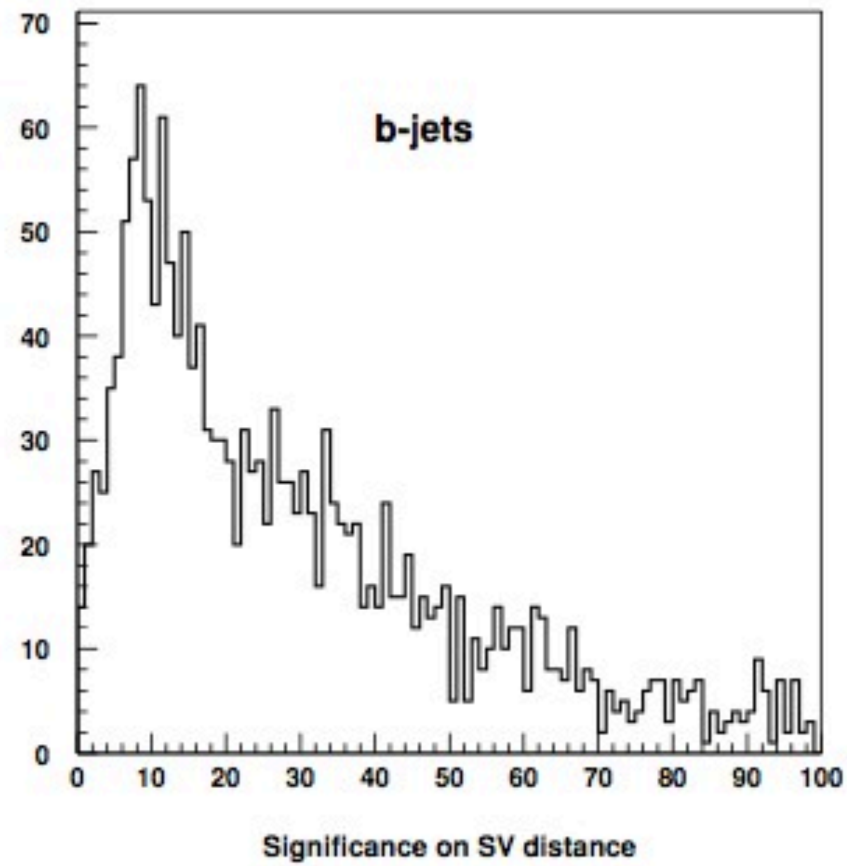
    - low resolution: 250 - 100 micometers for low (2-6) number of tracks (pT > 0.5 GeV)

    - high resolution: < 50 micometers: for more than 10 tracks associated to the vertex



SSV - Decay length





b-tagging Efficiency

- b-tagging algorithms are applied to all jets: heavy and light flavor
- measure the efficiency of the b-jets identification (**b-tag efficiency**) and light-jets (udsg) misidentification (**mis-tag rate**)
- b-tagging efficiency measurement algorithms: **pTrel**, **System8**

$$\vec{p}_T^{\,rel} = \frac{\vec{p}_\mu \times \vec{p}_{jet}}{p_{jet}}$$

- the momentum of the muon transverse to the jet axis

- larger for muons from b-hadrons than for muons from light flavor jets

Tagging two jets events

- Assume a sample of 100 events, each having exactly 2 jets
- Tagging efficiency in data = 0.6
- Tagging efficiency in MC = 0.7
- Scale factor (tagged jets) = 0.6/0.7 = 0.857
- Scale factor (non-tagged jets) = (1-0.6)/(1-0.7) = 1.33

$$weight_{event} = \prod_{Ntaggedjets} SF_{taggedjet} * \prod_{Nnon-taggedjets} \bar{SF}_{non-taggedjet}$$

Jets	MC uncorrected	Data	MC corrected
jet1 - tagged jet2 - untagged N events	70% 30% 21 events	60% 40% 24 events	SF = 0.857*1.33 23.94 events
jet1 - untagged jet2 - tagged Nevents	30% 70% 21 events	40% 60% 24 events	SF = 1.33*0.857 23.94 events
jet1 - tagged jet2 - tagged Nevents	70% 70% 49 events	60% 60% 36 events	SF = 0.875*0.875 36.00 events
jet1 - untagged jet2 - untagged Nevents	30% 30% 9 events	40% 40% 16 events	SF = 1.33*1.33 15.92 events

Correcting for tagging efficiencies Data/MC(1)

- Both Data and MC are tagged directly
- The tagging efficiency in MC is different from the tagging efficiency in data, for both heavy flavor and light jets
- MC needs to be corrected for tagging, using the scale factors:

  - for tagged jets : $SF_x = \frac{\varepsilon_x^{Data}}{\varepsilon_x^{MC}}$

  - for non-tagged jets : $\bar{SF}_x = \frac{1-\varepsilon_x^{Data}}{1-\varepsilon_x^{MC}}$

$$weight_{event} = \prod_{Ntaggedjets} SF_{taggedjet} * \prod_{Nnon-taggedjets} \bar{SF}_{non-taggedjet}$$

- This can be easily proved if we consider N events, each with n jets with the same flavor. The probability for k jets to be tagged is:

$$P_{n,MC}^{(k)} = C_n^k * \prod_k \epsilon_k^{MC} * \prod_{n-k} (1-\epsilon_k^{MC})$$

$$P_{n,Data}^{(k)} = C_n^k * \prod_k \epsilon_k^{Data} * \prod_{n-k} (1-\epsilon_k^{Data})$$

- The MC is corrected to data using the weight α_k

$$P_{n,Data}^{(k)} = \alpha_k * P_{n,MC}^{(k)}$$

Correcting for tagging efficiencies Data/MC(2)

- Involving α_k, this will mean:

$$C_n^k * \prod_k \epsilon_k^{Data} * \prod_{n-k} (1 - \epsilon_k^{Data}) = \alpha_k * C_n^k * \prod_k \epsilon_k^{MC} * \prod_{n-k} (1 - \epsilon_k^{MC})$$

$$\alpha_k = \prod_k \frac{\epsilon_k^{Data}}{\epsilon_k^{MC}} * \prod_{n-k} \frac{(1 - \epsilon_k^{Data})}{(1 - \epsilon_k^{MC})}$$

- The formula can be further generalized to the case of N events with n different jets.